

A model for setting services on auxiliary bus lines under congestion

E. Codina · A. Marín · F. López

Abstract In this paper, a mathematical programming model and a heuristically derived solution is described to assist with the efficient planning of services for a set of auxiliary bus lines (a bus-bridging system) during disruptions of metro and rapid transit lines. The model can be considered static and takes into account the average flows of passengers over a given period of time (i.e., the peak morning traffic hour). Auxiliary bus services must accommodate very high demand levels, and the model presented is able to take into account the operation of a bus-bridging system under congested conditions. A general analysis of the congestion in public transportation lines is presented, and the results are applied to the design of a bus-bridging system. A nonlinear integer mathematical programming model and a suitable approximation of this model are then formulated. This approximated model can be solved by a heuristic procedure that has been shown to be computationally viable. The output of the model is as follows: (a) the number of bus units to assign to each of the candidate lines of the bus-bridging system; (b) the routes to be followed by users passengers of each of the origin–destination pairs; (c) the operational conditions of the components of the bus-bridging system, including the passenger load of each of the line segments, the degree of saturation of the bus stops relative to their bus input flows, the bus service times at bus stops and the passenger waiting times at bus stops. The model is able to take into account bounds with regard to the maximum number of passengers waiting at bus stops and the space available at bus stops for the queueing of bus units. This paper demonstrates the applicability of the model with two realistic test cases: a railway corridor in Madrid and a metro line in Barcelona.

E. Codina (✉) · F. López

Departament d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya, Campus Nord, Building C5, C/ Jordi Girona, 31, 08034 Barcelona, Spain
e-mail: esteve.codina@upc.edu

A. Marín

E.T.S.I.A. Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Madrid, Plaza Cardenal Cisneros, 3, 28040 Madrid, Spain

Keywords Public transportation · Frequency design · Congestion modeling · Auxiliary bus lines

Mathematics Subject Classification (2000) 90B06 · 90B20 · 90B99 · 90C59 · 90C90

1 Introduction

The disruption of services in rapid transit and metro systems in large urban areas may have a considerable affect on the demand for mass transit services, particularly in cases of home-to-work and other mandatory trips. Many users of mass transit will switch their mode of transportation from their point of origin, resulting in an increase in car usage and road traffic. One means of reducing this problem is to bridge the stations at disrupted sections of the regular service (either rapid transit or metro) with auxiliary bus services that are able to operate under high-demand conditions.

These considerations, and the state of the art that will be described below, have motivated the present analysis and the development of a general model for setting establishing the frequencies for public transportation lines under congested conditions that can be applied to auxiliary bus services in response to large disruptions of the regular rapid transit and metro systems. In the test cases presented, the components of this model apply to these scenarios of disruption, and with the appropriate adaptations, the model could be adjusted to other public transportation systems.

These disruptions include lines that are completely interrupted or periods during which a number of transit stations are out of service inside an urban network for several hours. The model presented has been motivated by two test cases described in this paper. The first case consists of an important railway corridor in Madrid with an interruption of service at a segment of four consecutive stations. The second case consists of an interrupted portion of a metro line in Barcelona affecting ten consecutive metro stations.

Because the passenger capacity of auxiliary buses will probably be unable to accommodate all of the disrupted demand, these bridging services can be expected to experience bottlenecks. Thus, another purpose of the model is to identify, based on its solutions, these bottlenecks, and their causes and to take into account existing limitations on the deployment of the auxiliary bus system, such as the number of available bus units, the available space for bus queues and the maximum number of passengers waiting at bus stops. To build the model, basic concepts of queueing theory have been used for the average passenger flows at bus stops and the flow of bus arrivals in a way that allows for the reproduction of the interaction between both passengers and buses in an optimization model.

Because the resulting frequency-setting model uses average passenger flows over a given period of time (i.e., peak morning hours), it can be considered as a “static” model. This model requires the following basic input:

- (a) A set of bus lines, in a urban environment, identified by their segments and bus stops, that are candidates for operation during a disruption of the rapid transit or metro system.

-
- (b) An origin-to-destination trip table for the pick-up and drop-off stations in the disrupted area.

This trip table would be, in fact, the traversal origin–destination matrix of potential trips between the disrupted stations. The model assigns the available bus units to candidate bus lines and determines the number of services to be carried out according to the characteristics of each bus line. Bus lines discarded by the model have no bus units assigned to them.

The model minimizes the total set-up and/operational costs of the bus-bridging service plus the economic value of the total user travel time. The model is formulated with mixed-integer nonlinear programming, and the resulting problem is non-convex. It is solved with a specially designed heuristic method with reasonable computational times as shown in the test cases. The use of other methods for solving mixed-integer nonlinear problems applied to this model lies beyond the scope of this paper, and the authors are preparing a separate manuscript dedicated to more extensive computational tests.

1.1 State of the art

The design of bus line services has always been described in the scientific literature based on regular services under conditions in which congestion of the bus network is at a moderate level. Frequency-setting models have been formulated using passenger transit assignment schemas based on the concept of strategy (see, for instance, Spiess and Florian (1989) for a seminal paper on strategy-based passenger transit assignment) or passenger assignment schemas based on time tables. Using a strategy-based assignment model with a static approach, the models for frequency setting on public transportation networks in Constantin and Florian (1995) and Noriega and Florian (2003) must be noted. In cases in which lines are under strict timetable constraints, the important assignment models include those implemented for commercial packages, such as INRO Consultants (2010), TransCAD (2011) or PTV Vision VISUM 11.5 (2011). Other models, such as Zhao and Zeng (2008), have been very recently developed, as have those in Haupt et al. (2009) and in Papola et al. (2009). Models for frequency setting with elastic demand have been recently developed (Yoo et al. 2010). These models have typically been solved using heuristic methods (meta-heuristics) (see, for instance, Fan and Machemehl (2004) and Han et al. (2011)). Although establishing auxiliary bus services in cases of the disruption of regular transportation lines has been a common practice, only recently, in Kepaptsoglou and Karlaftis (2008), has this problem been directly taken into consideration. However, the paper in Kepaptsoglou and Karlaftis (2008) primarily addresses the design of a decision-support system for a bridging service. None of the papers cited here considers in detail the effects of congestion on the components of the transportation system (bus stops, bus queues, line cycles and passenger queues), takes into account external limitations or describes a method for the detection of bottlenecks.

1.2 Main contributions of the model

The main distinctive characteristics of the presented model for setting frequencies on transit lines can be summarized as follows:

-
- (a) The model takes into account the main factors involved in the congestion of transit lines using an analysis that can be extrapolated with small adaptations to situations other than bus-bridging systems.
 - (b) The model enables the detection of bottlenecks along transit lines as well as external limitations such as the space available for the queueing of units and the maximum number of passengers waiting at line stops.

With regard to the congestion modeling it must be pointed out that, under a conceptual point of view:

- (a) It considers the delay experienced by passengers at heavily congested transit stops resulting from the decrease of effective frequencies. This delay occurs when not all passengers can ride on the first-arriving unit because of irregular headway and the finite capacity of transit units. To address this issue, simulations of bus stops have been performed to determine the waiting time of passengers versus bus stop load factor relationships. These dynamics have been studied using bulk-service queue simulation models, and some of the results are shown in Appendix. Because the model is static, estimations of delays for passengers are based on average passenger flows over a given period of time, and Poissonian passenger arrivals at stops have been assumed. This approach is acceptable for modeling bus-bridging systems when the arrival of passengers at disrupted stations is based on frequent bulks. This frequency can occur in metro systems or at large railway stations with frequent train arrivals in which platoons of passengers are affected by a dispersion effect throughout the station. These conditions are met in both test cases presented in this paper. In situations in which the times between train arrivals are longer, a dynamic model considering arrival timetables would be more appropriate.
- (b) The queueing of transit units at stops is also modeled using approximated queueing models, and the effects of these queueing units on line cycles is also considered. Because the model is applied to bus lines in this case, passenger service times at stops are considered specifically for the case of bus systems. The relationship of passenger service times for bus units with the number of boarding and alighting passengers has been previously described with empirically developed formulas such as those in Guenther and Sinha (1983), Fernández and Tayler (2005), Jaiswal et al. (2008) and Fernández (2011).

In addition, the constraining of the inflow of buses at stops by limited bus stop capacities is considered using relationships described in TCRP Report 100 (2003). When applying the model to the design of a bus-bridging system, there will be a critical need for high-capacity bus stations, and the possibility of multiple-berth bus stops has been considered. The efficiency of multiple-berth bus stops has been taken into account following the empirical relationships described in TCRP Web Document 6 (1999). The model also considers the finite capacity available for accommodating waiting passengers at bus stops and the finite space for bus queues at stops.

Passengers are assigned to the bus lines of the bridging system using a congested shortest path criterion, i.e., all passengers of a given origin–destination pair are assigned to a unique path, which may consist of using several bus lines, and possibly, transfers between neighboring bus stops. This criterion was considered because of

its simplicity in providing direct information to passengers suffering the disruption. Passenger assignment models based on strategies (as in Spiess and Florian 1989) have not been considered in this paper based on the assumption that the crowding at bus stops would make it difficult to follow this behavior of passengers; however, the authors are considering this option for further research.

The main operational parameters for the bus lines resulting in a positive number of services are calculated as direct outcomes of the model: the number of services on each line, the bus service times at bus stops, estimations of bus units queue lengths and delays at bus stops, line cycles, estimations of passenger queue lengths and the riding/alighting volumes of passengers. The model also allows for the detection of bottlenecks in the bus system, making it possible to modify the system's configuration to enhance performance. These capabilities are illustrated with two realistic test networks in Sect. 8.

Computationally, the model has been formulated as a mixed-integer nonlinear programming problem, and a heuristic algorithm has specifically been developed to solve it. The heuristic algorithm consists of partially freezing some of the nonlinearities of the formulation and setting their values with an iterative procedure of the fixed-point type based on the method of successive averages (MSA).

Finally, because the presented model is a consequence of a general analysis of congestion in transit lines, it can be applied to systems other than bus-bridging lines, such as metro systems, with some adaptations. This is discussed in Sect. 5 in greater detail after the model formulation.

1.3 Layout of the paper

The layout of the paper is as follows. In Sect. 2 the sequence of operations and the times required from the point of view of buses and passengers are enumerated. In Sect. 3 the multicommodity network flow used to allocate passenger flows during the time horizon considered is described, as well as the set of definitions necessary to formulate the model in subsequent subsections. On this network flow it is possible to allocate boarding and alighting flows as well as transfer flows from line to line and passenger flows not using the auxiliary bus system. The transportation mode used for movements which do not use buses it is assumed to be pedestrian. Section 4 is dedicated to describe the modelling made of the operations described in Sect. 3: relationships assumed for passenger service times at bus stops, formulas for line cycle times, constraints expressing the maximum input flow of buses at bus stops, passenger waiting time at bus stops and link travel time functions. All these components appear together in Sect. 5 where a nonlinear mixed-integer programming problem is formulated. In Sect. 6 an approximation of the conceptual model described in previous Sect. 5 is made. This approximation consists of expressing the total waiting time of passengers at bus stops by means of a set of constraints and the addition of auxiliary variables. Section 7 describes the heuristic method used to find solutions to the model. Section 8 contains the description of the application of this model to two test networks: a railway corridor in Madrid comprising stations of Atocha, Recoletos, Nuevos Ministerios and Chamartín ("Tunel de la Risa's" corridor) and a set of metro stations in Barcelona's line 1 (from station Plaça d'Espanya to station Clot). A set of conclusions, possible extensions and possibilities of further research closes the paper.

2 Operational times in bus lines

In this paper, when referring to a bus-bridging system, a facility at which the bus units of the bridging system may perform the servicing of passengers will be called a bus stop, regardless of how complex such a facility may be. We will use the term “station” for the stations of the disrupted service that must be bridged. Bus stops of the bridging bus lines may be close enough to disrupted stations to make it unnecessary to distinguish between them in the model; in contrast, because of the characteristics of the street network topology, both bus stops and disrupted stations must be considered separately.

In the mathematical programming formulation described in this paper, for the assignment of services in the bus-bridging problem, a modeling is made of the different operational times involved in the usual operation of the bus lines. References Fernández and Tayler (2005) and Anderson and Scalia-Tomba (1981) provide a detailed description of these steps. Regardless of how complex a bus stop may be, the operational times described below conceive bus stops to be configured as depicted in Fig. 3 in Sect. 4.1. Very general configurations of bus stops (which are, in fact, bus stations) can be found in references TCRP Report 100 (2003) and TCRP Report 19 (1996).

Based on Fernández and Tayler (2005), the operational times to be considered are as follows:

Operational times V for vehicles

1. The time required for free movement between consecutive bus stops.
2. The maneuvering time required for entering a bus stop.
3. The time stationed at a bus stop.
 - (a) The queueing time required to access the boarding/alighting berth.
 - (b) The maneuvering time required for entering the boarding/alighting berth from the top of the queue (b1) and the clearance time or maneuvering time required to clear the boarding/alighting berth (b2).
 - (c) The passenger service time (PST) required for boarding/alighting (including the opening/closing of doors).
 - (d) The blocking time at berths: after servicing passengers, it may happen that the preceding bus cannot leave the bus stop, thus blocking the bus that has finished passenger service at the berth.
4. The time required for exiting the bus stop. This represents the elapsed period between leaving the berth and exiting the bus stop. This period of time is composed of
 - (a) Getting in the exit queue clearing the berth.
 - (b) Waiting in the exit queue.
 - (c) Waiting for an opportunity to incorporate into the external traffic flow, i.e., the reentry time.

Operational times P for passengers

1. Passengers who do not alight from the bus at the bus stop. Their times are the vehicle times $V1 + V2 + V3 + V4$.

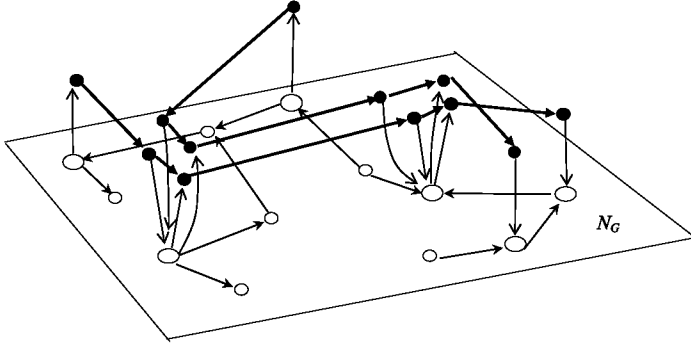


Fig. 1 Network representation by means of the graph $G = (N, A)$

2. Passengers who alight at the bus stop. Their times are $V1 + V2 + V3(b1)$ (berth access only) + $V3(a)$; the egress time from the vehicle must be taken also into account.
3. Passengers boarding a bus at a bus stop. Their times include the following:
 - (a) The time in the passengers' queue spent waiting for a bus of the appropriate line.
 - (b) The riding time + the time on board waiting for the bus with the doors closed. Note that this time is $V3(c)$.
 - (c) The time on board exiting the bus stop, which is equal to $V3(d) + V3(b2) + V4$.
 - (d) The elapsed time between the instant of reentry into the external traffic flow and the next bus stop, which is equal to time $V1$.

3 Network model

A basic component of this model is the network on which passenger movements take place. The network has a special structure to reflect these movements because of their relevance in congestion modeling. The model must capture passenger flows during a predefined period of time, including which passengers board at a stop to a given line or alight from a line to a stop, and which parts of the journey are made by bus or by an alternate transportation mode. For simplicity, the alternate mode is considered to be walking.

Passenger flows go through a directed graph $G = (N, A)$, whose structure is sketched in Fig. 1. This graph will be referred to in the following as the expanded transit network, the expanded graph or the expanded bus network. The set of nodes N splits into two subsets, N_G and $N \setminus N_G$. Nodes in the set N_G and links $a = (i, j)$, so that $i \in N_G$ and $j \in N_G$, are used to model transfer movements or simply trips balking at the bus transportation system and accomplished using the alternative transportation mode. The set of such links will be designated by A_G :

$$A_G = \{(i, j) \in A \mid i \in N_G, j \in N_G\} \quad (1)$$

On the graph, incoming and outgoing nodes from a given one $i \in N$ are designated by $E(i)$, $I(i)$, respectively. A subset of nodes within N_G , those that will be considered as representing disrupted stations or bus stops, will be denoted by \hat{N}_G . L designates the set of bus lines and $\Pi_\ell = \{b_1, \dots, b_{\mu^\ell}\}$ is the ordered set of μ^ℓ bus stops or stations $b_j \in \hat{N}_G$ on line $\ell \in L$. Conversely, given a bus stop $b \in \hat{N}_G$, L_b will designate the set of bus lines containing stop b . The subset of links A_b is composed of links $a = (i, j)$ so that neither $i \in N_G$ nor $j \in N_G$ and a is an entering link to stop b :

$$A_b = \{(i, j) \in A \mid i, j \notin N_G, \exists (j, k) \in E(j) \text{ so that } k \in I(b)\} \quad (2)$$

A_ℓ is the set of links modeling line $\ell \in L$ so that $i, j \notin N_G$.

$$A_\ell = \{(i, j) \in A \mid i, j \notin N_G, (i, j) \text{ belongs to line } \ell\} \quad (3)$$

$A_{b,\ell}$ defines the set of links $a \in A_b$ corresponding to line $\ell \in L$.

$$A_{b,\ell} = A_b \cap A_\ell \quad (4)$$

The set of lines incoming to a bus stop b can then be defined by $L_b = \{\ell \in L \mid A_{b,\ell} \neq \emptyset\}$. The following definitions are also useful:

$$A_\ell^x = A_\ell \setminus \bigcup_{b \in \Pi_\ell} A_{b,\ell} \quad (5)$$

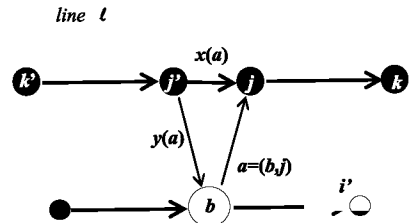
$$A_\ell^0 = A_\ell \setminus A_\ell^x \quad (6)$$

At a node $i \in N$, $E(i)$ will designate the forward star of outgoing links emerging from i . $\hat{E}(i)$ will designate the outgoing links (i, j) so that $j \notin N_G$:

$$\hat{E}(i) = \begin{cases} E(i) & \text{if } i \notin N_G \\ E(i) \setminus (E(i) \cap A_G) & \text{if } i \in \hat{N}_G \\ \emptyset & \text{if } i \in N_G \setminus \hat{N}_G \end{cases} \quad i \in N \quad (7)$$

Figure 2 shows in greater detail the set of links modeling alighting and boarding operations at bus stops for a given line $\ell \in L$. Links $a = (b, j)$, $b \in \Pi_\ell$, $j \notin N_G$ capture boarding and waiting time at a bus stop for passengers willing to board on buses of line ℓ . For a given bus stop b belonging to line ℓ , the boarding link from $b \in \hat{N}_G$ to line ℓ will be designated by $a(\ell, b)$ if it exists, and, with the help of the

Fig. 2 Detail of the expanded graph for modeling bus lines and stops. $b, i' \in N_G$, $(k', j') \in A_b$. $x(a) = (j', j)$, $(j, k) \in A_\ell$. Nodes j, j' for boarding and alighting, respectively, from/at node $b \in \Pi_\ell$ for line ℓ are given by $j = j_\ell(b)$, $j' = j'_\ell(b)$



operator $j_\ell(b)$, by $a = (b, j_\ell(b))$. For a boarding link $a = (b, j_\ell(b))$, link $x(a)$ if it exists, denotes the link on which passengers wait on board the bus without alighting at that bus stop, whereas $y(a)$ represents the corresponding alighting link, if it exists. In addition, an alighting link at a bus stop $b \in \hat{N}_G$ for a line $\ell \in L$ will be designated by their head and tail nodes $(j'_\ell(b), b)$ with the help of operator $j'_\ell(b)$.

Passenger demand is represented by a set of positive flows, $g_\omega > 0$, each for a given origin–destination pair $\omega = (p, q)$, during a predefined period of time of length H units. The set of origin–destination pairs W is defined from nodes within N_G , that are disrupted stations. D is the set of nodes that are destinations. The origin–destination trip matrix will be designated by g_ω , $\omega \in W$. Indices ω , or explicitly (p, q) , will be used for an origin–destination pair when considered convenient.

Passenger flows during the specified period of time will be organized in commodities, one per origin–destination, and the total flow of passengers on a link $a \in A$, during that period of time and for an origin–destination pair $\omega \in W$ will be denoted by v_a^ω . The balance equations for flows per origin–destination pair $\omega = (p, q) \in W$ at a node $i \in N$ will be as follows:

$$\sum_{j \in E(i)} v_{ij}^\omega - \sum_{k \in I(i)} v_{ki}^\omega = \begin{cases} -g_{(p,q)} & \text{if } i = q \\ g_{(p,q)} & \text{if } i = p \\ 0 & \text{if } i \neq p \end{cases} \quad (8)$$

By adding non-negativity conditions $v_a^\omega \geq 0$ for flows on links to previous relationships (8), the following polyhedra are defined:

$$V^\omega = \{v^\omega \in R_+^{|A|} \mid v^\omega = (\dots, v_{i,j}^\omega, \dots; (i, j) \in A), v_{i,j}^\omega \text{ verify (8)}\}, \quad \omega \in W \quad (9)$$

Finally, the polyhedron \mathcal{V} of total passenger flows on links will be

$$\mathcal{V} = \left\{ v \in R_+^{|A|} \mid v = \sum_{\omega \in W} v^\omega, v^\omega \in V^\omega \right\} \quad (10)$$

4 Modeling operational times

It will be assumed that there are p bus units available to be assigned to $|L|$ predetermined candidate bus lines, with all bus units of equal type and with capacity for c passengers. Thus, if $n^\ell \geq 0$, $n^\ell \in \mathbb{Z}$, $\ell \in L$ is the number of units assigned to line $\ell \in L$, then

$$\sum_{\ell \in L} n^\ell \leq p \quad (11)$$

4.1 Modeling bus stops

The performance of bus stops plays a fundamental role in the performance of bus transport. Empirically, bus stops have been extensively studied under various configurations and operational conditions. Reference TCRP Report 100 (2003) provides clear descriptions of these aspects in situations in which the probability of queuing

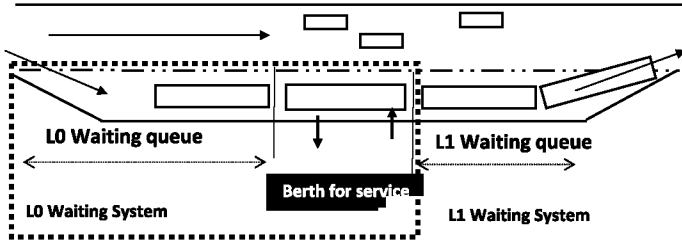


Fig. 3 A schematic representation for a bus stop

of buses at stops is low or moderate. In this paper, it will be assumed that this probability may be high, and to this end, the operation of buses at stops will be modeled using two waiting systems in tandem with blocking as shown in Fig. 3. The entering and exiting waiting systems will be referred to as system L0 and system L1, respectively. Although bus stops may have a physical configuration other than the bay-type depicted in Fig. 3, conceptually, these two waiting systems can always be considered. In addition, because bus stops of the bus-bridging system should be as close as possible to the disrupted stations of the public transportation system, linear on-street or bay-type bus stops will be assumed. Linear bus stops are the most common configuration that will be possible to implement because only very relevant stations of the disrupted transport system may have facilities that could permit better configurations of the bus stops of the auxiliary system. Linear bus stops are characterized by having berths one behind the other so that when the bus stop is empty, the first bus to arrive occupies the first loading area, the second bus occupies the second loading area, and so on. See TCRP Report 100 (2003) for a discussion of the effectiveness of linearly configured bus stops. Waiting system L1 may block waiting system L0, i.e., a bus having finished its dwell time at a berth may not proceed because the L1 space is still full. In this paper, it will be assumed that the mean time w_b^1 that a bus stays in space L1 of bus stop $b \in N_G$ is fixed. This quantity is assumed to include the clearance time from the berth and the reentry time of buses into the external traffic flow (times V4(a) + V4(b) + V4(c) in Sect. 2). The clearance time is more or less linear with bus length, and, following TCRP Report 100 (2003), a value of 1 second per meter will be adopted. Also, following TCRP Report 100 (2003), reentry time is assumed to be linear with external traffic flow in the range 0–700 veh/h at a rate of 1 second every 100 veh/h. A final assumption is that bus stops are located far enough from traffic signals to decrease bus service time.

4.1.1 Modeling bus service time and passenger service time at bus stops

Assuming for simplicity that bus stops are served only by a single bus line, the mean service time κ_b at a bus stop $b \in \hat{N}_G$ will be expressed in three terms:

$$\kappa_b = \kappa_b^M + \kappa_b^{OM} + \kappa_b^{PST}, \quad b \in \hat{N}_G \quad (12)$$

where κ_b^M is the maneuver time V3(b) for a bus to reach the boarding/alighting berth from the head of the bus queue in space L0, κ_b^{OM} is an operational margin and κ_b^{PST} is

the mean passenger service time V3(c) for a bus as a result of the boarding/alighting of passengers at the stop's berth. κ_b^M depends basically on the bus stop configuration and may be assumed to be constant.

Passenger service times have been studied by several authors, mainly from an empirical point of view. The reader may consult references Guenther and Sinha (1983), Fernández and Tayler (2005), Jaiswal et al. (2008) or Fernández (2011) for empirical expressions of the average passenger service time. Following Fernández and Tayler (2005), for instance, the passenger service time for parallel boarding/alighting operations can be summarized with a general formula such as this:

$$\kappa_b^{PST} = a_b + \max_{1 \leq i \leq n_+, 1 \leq j \leq n_-} \{\tau_{b,i}^+, \tau_{b,j}^-\}, \quad b \in \hat{N}_G \quad (13)$$

where n_+ and n_- are the number of doors used as entrances and exits of the bus, respectively, and $\tau_{b,i}^+$ and $\tau_{b,j}^-$ are the times necessary for embarking/disembarking the corresponding volumes of passengers per service. a_b can be considered a dead time resulting from doors opening/closing. The relationship between these times and the number of passengers $u_{b,i}^+, u_{b,j}^-$ per bus entering/exiting through the doors has been examined in a number of papers under a variety of specific conditions (see Fernández and Tayler 2005 or TCRP Report 100 2003). The most accepted assumption is that this relationship is linear:

$$\tau_{b,i}^+ = g^+ u_{b,i}^+, \quad \tau_{b,j}^- = g^- u_{b,j}^-, \quad b \in \hat{N}_G, \quad 1 \leq i \leq n_+, \quad 1 \leq j \leq n_- \quad (14)$$

where g^+ and g^- are the boarding and alighting service time per passenger, respectively. These times depend on the number of doors used for boarding/alighting and on whether these operations are carried out using single-channel or multiple-channel facilities and on whether the boarding/alighting operations are carried out in serial or in parallel. Thus $g^+ = g^+(n^+)$, $g^- = g^-(n^-)$. It will be assumed that bus units have separate doors for boarding/alighting operations in parallel (two rear doors for alighting and one door for boarding).

Assuming now the general case, in which a bus stop operates for several bus lines, it is possible to express the average bus service time for the set of lines stopping at stop b as a function of the total passenger volumes v on the expanded network links and the number of services z^ℓ for line $\ell \in L_b$ entering at stop $b \in N_G$. The specific passenger service time $\kappa_{b,\ell}^{PST}(v, z)$ for buses of line $\ell \in L$ at a stop $b \in \Pi_\ell$, will be given by

$$\kappa_{b,\ell}^{PST}(v, z) = \max \left\{ g^+(n^+) \frac{v_a}{z^\ell}, g^-(n^-) \frac{v_{y(a)}}{z^\ell} \right\}, \quad b \in \Pi_\ell, \ell \in L \quad (15)$$

In cases of linear bus stops (on-line or off-line) in which the berth's length allows for more than one loading area where two or more buses can be served simultaneously, an efficiency factor e for the berth must be used, which depends on the maximum number of loading areas s_b that are possible to allocate and on whether arrival of buses is at random or it is forced to be platooned. The efficiency factor for linear bus stops varies for on-line and off-line configurations. These efficiency factors will be adopted following TCRP Report 100 (2003). Accordingly, the average effective

bus service time (times V3(b) + V3(c) as described in Sect. 2) for units of line $\ell \in L$ at bus stop $b \in \Pi_\ell$ will be given by

$$\kappa_{b,\ell}(v, z) = \frac{\kappa_b^M + \kappa_b^{OM} + \kappa_{b,\ell}^{PST}(v, z)}{e(s_b)}, \quad b \in \Pi_\ell, \ell \in L \quad (16)$$

From this, the average passenger service time of buses entering at bus stop b will be given by

$$\kappa_b^{PST}(v, z) = \frac{\sum_{\ell \in L_b} z^\ell \kappa_{b,\ell}^{PST}(v, z)}{\sum_{\ell \in L_b} z^\ell}, \quad b \in \hat{N}_G \quad (17)$$

Accounting for the efficiency factor, the average effective bus service time per physical loading area at bus stop $b \in \hat{N}_G$ will be given by

$$\kappa_b(v, z) = \frac{\kappa_b^M + \kappa_b^{OM} + \kappa_b^{PST}(v, z)}{e(s_b)}, \quad b \in \hat{N}_G \quad (18)$$

4.1.2 Line cycles and bus-queueing at bus stops

Assuming that n^ℓ buses have been assigned to line ℓ out of the total p buses available and assume that they perform a total of z^ℓ runs on line $\ell \in L$, the load factor ρ_b at stop b for the input bus-waiting system L0 will be expressed by the ratio between the bus arrivals per unit of time and the berth's effective maximum capacity or the berth's maximum number of services per unit time:

$$\rho_b = \frac{\text{Arrival of buses per unit of time}}{\text{Maximum number of services per unit of time}}, \quad b \in \hat{N}_G \quad (19)$$

The arrival of buses per unit of time at stop b during a period of H units of time must be $\frac{1}{H} \sum_{\ell \in L_b} z^\ell$; however, the maximum number of services per unit of time will be s_b/κ_b . The loading factor will then be expressed as

$$\rho_b(v, z) = \frac{\kappa_b(v, z)}{s_b H} \sum_{\ell \in L_b} z^\ell, \quad b \in \hat{N}_G \quad (20)$$

The waiting times at bus queues w_b^0 at stop $b \in N_G$ can be expressed directly as a function $W_b(\cdot)$ of this traffic factor using Allen–Cunneen's approximation (see, for instance Allen 1998), which provides an approximation of the queueing delay for general $GI/G/s$ queues. This leads to $w_b^0 = W_b(\rho_b)$. Because of this, w_b^0 can also be considered as a function of (v, z)

$$w_b^0(v, z) = W_b(\rho_b(v, z)), \quad b \in \hat{N}_G \quad (21)$$

The waiting times of buses at waiting system L1 at a stop $b \in N_G$, w_b^1 , will be assumed as fixed. Because the queueing of buses at stops and passenger service times are also dependent on (v, z) , bus cycles C_ℓ on line $\ell \in L$ are also dependent on (v, z) :

$$C_\ell(v, z) = C_\ell^0 + \sum_{b \in \Pi_\ell} (\kappa_{b,\ell}(v, z) + w_b^1 + w_b^0(v, z)), \quad \ell \in L \quad (22)$$

where C^0 is the result of the free travel time from stop to stop. This free travel time will be represented by t_a^0 at a link $a \in A_\ell^0$. This link $a = (j_\ell(b), j'_\ell(b'))$ goes from stop b to stop b' . Therefore, t_a^0 is the elapsed time beginning immediately after leaving stop b and upon arrival at the end of the (possibly non-empty) queue of buses in b' , and, thus,

$$C_\ell^0 = \sum_{a \in A_\ell^0} t_a^0, \quad \ell \in L \quad (23)$$

Another assumption is that the fixed dead times of buses used for changing of drivers etc. during the time horizon H are made at depots located outside any of the bus stops in the network and that buses are then withdrawn from the line during these dead times. This makes the effective operational time horizon for buses assigned to line $\ell \in L$, H^ℓ smaller than or equal to H . The relationship between the number of services z^ℓ and the number of buses n^ℓ assigned to a line $\ell \in L$ is then

$$n^\ell H^\ell \geq z^\ell C_\ell(v, z), \quad \ell \in L \quad (24)$$

4.1.3 Maximum input flow of buses at a bus stop

The maximum input flow of buses that a bus may be able to serve is determined primarily by two factors. The first factor is the bus service time, which depends on the flow of passengers arriving at the bus stop and the time required by buses to approach the service berth plus the clearance time and (possibly) the blocking time that a loaded bus unit ready to leave may encounter as a result of other buses trying to reenter in the external traffic flow. The second factor is a physical limitation on the space in which buses may queue to get to the service berth of the bus stop. Clearly, the greater the input flow of buses to the bus stop, the longer the waiting queues of buses, potentially causing the waiting units to block external facilities such as intersections or pedestrian crosswalks. Because of that spillback phenomenon, in this model, a limitation on the maximum length of the bus queues will be imposed indirectly. The model actually limits the maximum input flow of buses that may enter a bus stop; this indirectly limits the average bus queue length at the bus stop, and, thus, the possibility of spillback is reduced. An estimation obtained by simulation of the ratio between the average queue length and the maximum queue length that is only surpassed 5% of the time will be used. Toward this end, let \mathcal{L}^0 , \mathcal{L}^1 , η^0 and η^1 be defined as follows:

- \mathcal{L}^0 : The maximum number of buses allowed in space 0 (i.e., those waiting to get to the boarding berths + the number of buses allowed in the berths).
- \mathcal{L}^1 : The maximum number of buses allowed at space 1 (i.e., those that can be allocated at the exit of the boarding berths and the reentry of buses into road traffic).
- η^0 and η^1 : The occupancy factors at 95% or the ratio between the average occupancy at spaces 0 and 1 and the maximum occupancy attained 95% of the time (usually between 3 and 4). A value of 3 is adopted.

Applying Little's formula to both waiting systems 0 and 1 at bus stop $b \in \hat{N}_G$:

$$\frac{(\kappa_b(v, z) + w_b^0(v, z))}{H} \sum_{\ell \in L_b} z^\ell \leq \frac{\mathcal{L}_b^0}{\eta^0}, \quad b \in \hat{N}_G \quad (25)$$

$$\frac{w_b^1}{H} \sum_{\ell \in L_b} z^\ell \leq \frac{\mathcal{L}^1}{\eta^1}, \quad b \in \hat{N}_G \quad (26)$$

In addition, if ρ_b is the loading factor for the bus stop, it is desirable that $\rho_b < 1$. Taking, for instance, $\epsilon = 0.02$ or 0.05 :

$$\frac{\kappa_b(v, z)}{s_b H} \sum_{\ell \in L_b} z^\ell = \rho_b \leq 1 - \epsilon, \quad b \in \hat{N}_G \quad (27)$$

In summary:

$$\sum_{\ell \in L_b} z^\ell \leq \hat{Z}_b(v, z) \triangleq H \min \left\{ \frac{(1 - \epsilon)s_b}{\kappa_b(v, z)}, \frac{\mathcal{L}^0}{\eta^0(\kappa_b(v, z) + w_b^0(v, z))}, \frac{\mathcal{L}^1}{\eta^1 w_b^1} \right\},$$

$$b \in \hat{N}_G \quad (28)$$

4.1.4 Limitations on the maximum admissible headway and frequencies of lines

The solutions of our model may be required to provide a number of services z^ℓ at line $\ell \in L$ so that frequencies on that line fall within a predetermined range of frequencies, say $[\underline{f}, \hat{f}]$, provided that $z^\ell > 0$. The lower bound \underline{f} for frequencies will be expressed in terms of a maximum headway h_{\max} , i.e. $\underline{f} = H/h_{\max}$. The following constraints would enforce these conditions:

$$z^\ell \geq \lambda_\ell \frac{H}{h_{\max}}, \quad z^\ell \leq \lambda^\ell H \hat{f}, \quad \lambda_\ell \in \{0, 1\}, \ell \in L \quad (29)$$

Clearly, if $\lambda^\ell = 1$, a positive number of services will be assigned to line ℓ , the resulting headway will be $H/z^\ell \geq h_{\max}$, and the frequency on that line will not exceed \hat{f} , whereas if $\lambda^\ell = 0$, then no services will be assigned to the line.

4.2 Total waiting time of passengers at a bus stop

The waiting time of passengers at bus stops for boarding a bus may not be a negligible component of the total trip time, particularly in congested situations. It is commonly accepted that passengers waiting at a bus stop can be modeled with a queuing process with the following characteristics:

1. Waiting passengers are in a bulk-service queue, i.e., passengers' arrivals are on a one-to-one basis, but there is batch-service for each bus arrival at the stop. Usually, the total time used for the bus service is shorter than mean bus interarrival time at the bus stop.
2. The number of passengers that can be allocated at each arriving bus is a random variable that is independent from the number of passengers waiting at the moment of arrival.

Bulk-service queues play a fundamental role in transportation and logistics, and have undergone extensive theoretical study. The analysis of such models was originated by Bailey (1954) for the case of a fixed capacity of servers. The case of a random capacity of servers was studied by Jaiswal (1964). A particularly well-known example of a bulking queue model is $M/M^{[Y]}/1$ (see, for instance, Jain et al. 2007). However, not even in this case can closed formulas be obtained for the waiting time, and in order to calculate the waiting time, it is necessary to obtain zeroes of a function in the complex plane. This is one reason these models have not been used in applications of passenger transportation planning. Powell (1986), obtained closed formulas for approximating the moments of waiting times and queue lengths for bulk-arrival and bulk-service queues, which are useful for applications at terminals of networks in freight transportation models, but which cannot be applied for passenger queues at bus stops, as the range of the parameters considered in Powell (1986) cannot be adapted to the case of public transportation networks. In TCRP Report 100 (2003) it is suggested that the waiting time of passengers at stops can be estimated using an $M/M/s$ queuing model. However, as shown in Appendix, this may lead to a strong underestimation of passenger waiting times at congested bus stops.

For the case of the $M/M^{[Y]}/1$ bulk-service model with random service capacity, the average waiting time per passenger φ for a given bus line at a stop is given as a function of the load factor ϱ of the queue of the form:

$$\varphi(\varrho) = P_0 \xi(\varrho) \quad (30)$$

where $\xi(\varrho)$ is an augmenting factor resulting from the congestion for which there is no closed formula, excluding some particular cases (i.e., constant service capacity), and P_0 is the waiting time at the bus stop per service and passenger in an uncongested situation. P_0 depends on the mean and deviation of the bus interarrival time, h and σ , respectively (see, for instance, Trivedi 2002):

$$P_0 = \frac{h}{2} \left(1 + \frac{\sigma^2}{h^2} \right) \quad (31)$$

For the estimation of variance σ^2 , the method of Adebisi (1986) or the approximation in Vasanthakumar (1993) can be used; however, to simplify our model, it will be assumed in advance that lines operate with short headways (< 10 minutes) and that headway adherence occurs frequently (level of service D or E). Following TCRP Report 100 (2003), the coefficient of variation $\chi = \sigma/h$ will be assumed fixed a priori and in the interval $[0.40, 0.74]$.

For the more general bulk-service queue $M/G^{[Y]}/1$ with random service capacity, (30) also holds, and the randomness of the capacity of the server on passenger delay is not very significant, as shown in Appendix.

For a given line $\ell \in L$ and bus stop $b \in \Pi_\ell$, let $a = a(\ell, b)$ be the corresponding link in the expanded network. If z^ℓ is the number of services on that line, the mean interarrival time will be H/z^ℓ . The load factor ϱ_a for the queue of passengers willing to board servers of line ℓ will be given as a function of link flows of the expanded network by

$$\varrho_a(v, z) = \frac{v_a}{cz^\ell - v_{x(a)}} a = a(\ell, b), \quad b \in \Pi_\ell, \ell \in L \quad (32)$$

where $v = \sum_{\omega \in W} v^\omega$. Thus, $\varrho_a(v, z)$ is the ratio between the average flow of boardings at line ℓ at bus stop b and the average capacity available for service during period H . Implicitly, it will be assumed that $\varrho_a < 1$. In ratio (32), cz^ℓ is the maximum number of passengers that can be alighted from bus stop b by z^ℓ services carried out by bus units with maximum capacity c .

The average queueing time $\varphi_a(v, z)$ per passenger at a bus stop waiting for buses, if the headway's coefficient of variation, χ_a , is fixed a priori, will be given by

$$\varphi_a(v, z) = \frac{H}{2z^\ell} (1 + \chi_a^2) \xi_a(\varrho_a(v, z)), \quad a = a(\ell, b) b \in \hat{N}_G, \ell \in L_b \quad (33)$$

Thus, in general, we will accept the following product form for the passenger's waiting time for line ℓ at a stop b :

$$\varphi_a(v, z) = P_a(z) \xi_a(\varrho_a(v, z)) \quad (34)$$

Accordingly, if $a = a(\ell, b)$, the total passenger waiting time for boarding at a line, ζ_a , will be expressed as a function of (v, z) as

$$\zeta_a(v, z) = v_a \varphi_a(v, z) = v_a P_a(z) \xi_a\left(\frac{v_a}{cz^\ell - v_{x(a)}}\right) \quad (35)$$

4.3 Modeling finite capacity for passengers waiting at a bus stop

The limited capacity of stops to accommodate queues of passengers waiting to board buses must also be taken into account. Assume that \hat{N}_b^{pax} is the maximum number of passengers that bus stop $b \in N_G$ is able to accommodate. It is assumed that there may be a queue of passengers for each of the lines stopping at the stop. The sum of the corresponding queue lengths might not exceed the capacity \hat{N}_b^{pax} of the stop, at least during a large fraction α of the time (say $\alpha = 0.95$). The application of Little's formula to the queue of passengers waiting for buses of line ℓ at bus stop $b \in \Pi_\ell$ leads to the following formula:

$$\frac{v_a}{H} \xi_a = \text{Average \# of pax. in queue associated to link } a = a(\ell, b), b \in \Pi_\ell, \ell \in L \quad (36)$$

Let η_b be the ratio between the queue length exceeded only a fraction $1 - \alpha$ of the time and the average queue length for bus stop $b \in \hat{N}_G$. The model requires that the sum of all of the queue lengths does not exceed a stop's capacity \hat{N}_b^{pax} during a fraction α of the time under the following constraints:

$$\sum_{\ell \in L_b} \xi_{a(\ell, b)}(v, z) \leq \frac{H}{\eta_b} \hat{N}_b^{\text{pax}}, \quad b \in \hat{N}_G \quad (37)$$

4.4 Link travel time functions on the expanded transit network

If $a = a(\ell, b)$ is a boarding link in the expanded graph at stop $b \in N_G$ for line $\ell \in L$, the corresponding operational times assigned to that link will P3a and V3c (for line ℓ),

i.e., the time from the passenger's arrival at the bus stop until the bus closes its doors to initiate clearance of the berth. Taking into account the efficiency factor of the bus stop, the link travel time $t_a(v, z)$ will be

$$\begin{aligned} t_a(v, z) &= \frac{\kappa_{b,\ell}^{PST}(v, z)}{e(s_b)} + \varphi_a(v, z) & \text{if } z^\ell > 0 \\ t_a(v, z) &= +\infty & \text{if } z^\ell = 0 \\ a &= a(b, \ell), b \in \Pi_\ell, \ell \in L \end{aligned} \quad (38)$$

The corresponding alighting link $y(a)$ for the previous boarding link $a = a(\ell, b)$, related both a and $y(a)$ to line $\ell \in \mathcal{L}$ and stop $b \in \Pi_\ell$, will be assigned time P2b, and this time will be assumed to be a linear function of the average number of alighting passengers per service:

$$\begin{aligned} t_{y(a)}(v, z) &= \frac{g^-(n^-)}{2} \frac{v_{y(a)}}{z^\ell} & \text{if } z^\ell > 0 \\ t_{y(a)}(v, z) &= \frac{g^-(n^-)}{2} c & \text{if } z^\ell = 0 \\ a &= a(b, \ell), b \in \Pi_\ell, \ell \in L \end{aligned} \quad (39)$$

For consecutive stops b, b' in line $\ell \in L$, let us consider link $a = (j_\ell(b), j'_\ell(b'))$. Time t_a for these links will be from exiting the berth of bus stop b until entering the berth of stop b' . Thus, they will be assigned times V1 + V2 + V3a + V3b1 for bus stop b' plus times V3b2 + V3d + V4a + V4b for bus stop b . In our model, t_a will be expressed as the sum of a constant term \tilde{t}_a^0 plus the bus queueing time V3a, $w_{b'}^0(v, z)$, at the next stop b' . This constant term is then composed of the times $t_a^0 = V1 + V2$ plus the berth's clearance time, the waiting time V4a, the reentry time V4b for exiting the bus stop b and the maneuvering time V3b1, $\kappa_{b'}^M$, at stop b' . These last three terms will be assumed constant in our model, i.e., $\tilde{t}_a^0 = t_a^0 + w_b^1 + \kappa_{b'}^M$. The blocking time V3d will be considered negligible as it will be assumed that the locations of bus stops and their configurations facilitate smooth reentry into the external traffic flow

$$t_a(v, z) = \tilde{t}_a^0 + w_{b'}^0(v, z), \quad a = (j_\ell(b), j'_\ell(b')), b, b' \in \Pi_\ell, \ell \in L \quad (40)$$

If $a = (j'_\ell(b), j_\ell(b))$ for line $\ell \in L$, then this link will be assigned a travel time function $t_a(v, z)$, which reflects time V3c for line ℓ . Taking into account the efficiency factor of the bus stop, the travel time function will be given by

$$t_a(v, z) = \frac{\kappa_{b,\ell}^{PST}(v, z)}{e(s_b)}, \quad a = (j'_\ell(b), j_\ell(b)), b \in \Pi_\ell, \ell \in L, \quad (41)$$

Finally, the travel times of the links on the pedestrian network will be assumed constant, i.e., $t_a(v, z) = t_a^P$, $a = (i, j)$, $i \in N_G$, $j \in N_G$.

5 The conceptual model expressed as a mathematical programming problem

With the definitions of Sect. 4.4, the total travel time spent by all passengers will be simply expressed as

$$\text{Total travel time} = \sum_{a \in A} v_a t_a(v, z),$$

and the objective function of the design model will consist of economical costs for fleet recruiting and service assignment plus the economical cost of the passengers total travel time using an estimation θ of the social cost of time (for instance, in €/min). In addition, the following definitions related to the link travel times defined in Sect. 4.4 will be given for convenience:

$$T_a(v, z) = \begin{cases} \frac{\kappa_{b,\ell}^{PST}(v, z)}{e(s_b)} & \text{if } a = (i, j), i \in N_G, j \notin N_G \\ t_a(v, z) & \text{otherwise} \end{cases} \quad (42)$$

With these definitions, the total travel time on a network given by $\sum_{a \in A} v_a t_a(v, z)$ will be expressed in terms of link travel times T_a in (42) and the total waiting times of passengers at bus stop ζ_a in (35) as follows:

$$\sum_{a \in A} v_a t_a(v, z) = \sum_{a \in A} v_a T_a(v, z) + \sum_{\ell \in L} \sum_{b \in \Pi_\ell} \zeta_{a(\ell, b)}(v, z) \quad (43)$$

Another component to be considered to definitively outline the setting of services is the passenger assignment to bus lines. An aspect directly related to this component is to provide information to users of the bridging system about the routes or lines that are most convenient for them to reach their destination. The assignment of services is conditioned by the assignment of passengers to bus lines and, for a given origin destination pair $\omega \in W$, an assignment of g_ω passengers may split this amount into several routes on the expanded transit network. This splitting of origin destination flows may be very difficult to implement in a realistic situation. In this model, it will be assumed that all the demand for an origin–destination is assigned to a single route on the expanded transit network, thereby providing the solution that is easier to put into practice from the perspective of the information to be provided to passengers.

The following model, M0, provides, as basic results, the number of bus units to be assigned to a bus a line along with the total number of services to be carried out by these units, taking into account the modeling of congestion described in previous Sect. 4. In addition, because of the information requirements to the passengers of the bridging system, the solutions of the model must be based on the use by the passengers of unique routes, for each origin destination pair. Toward this end, additional constraints R1 and R2 in model M0 are added to enforce this condition on passenger

flows on the expanded transit network.

$$\begin{aligned}
\text{[Model M0]} \quad & \text{Min}_{n,z,v,\tau,\lambda} \sum_{\ell \in L} (\varsigma^\ell n^\ell + \gamma^\ell z^\ell) + \theta \sum_{a \in A} v_a T_a(v, z) + \theta \sum_{\ell \in L} \sum_{b \in \Pi_\ell} \zeta_{a(\ell, b)}(v, z) \\
\text{A1} \quad & \text{s.t.:} \quad \sum_{\ell \in L} n^\ell \leq p \\
\text{A2} \quad & H^\ell n^\ell \geq z^\ell C_\ell(v, z), \quad \ell \in L \\
\text{A3} \quad & 0 \leq z^\ell \leq \lambda^\ell \hat{f} \cdot H, \\
\text{A4} \quad & z_\ell \geq \lambda^\ell \frac{H}{h_{\max}}, \quad \ell \in L \\
\text{B1} \quad & v \in \mathcal{V} \\
\text{R1} \quad & \sum_{a \in E(i)} \tau_a^\omega \leq 1, \quad i \in N, \omega \in W \\
\text{R2} \quad & v_a^\omega \leq M \tau_a^\omega, \quad a \in A, \omega \in W \\
\text{Qb1} \quad & \sum_{\ell \in L_b} z^\ell \leq \hat{Z}_b(v, z), \quad b \in \hat{N}_G \\
\text{Qp1} \quad & v_a + v_{x(a)} \leq c z^\ell \\
\text{Qp2} \quad & \sum_{\ell \in L_b} \zeta_a(v, z) \leq \frac{H}{\eta_b} \hat{N}_b^{\text{pax}} \\
& a = a(\ell, b), b \in \Pi_\ell, \ell \in L \text{ in Qp1, Qp2} \\
\text{Var. Def.} \quad & n^\ell \geq 0, \quad n^\ell \in \mathbb{Z}, \ell \in L \\
& z^\ell \in \mathbb{Z}, \quad \lambda_\ell \in \{0, 1\}, \ell \in L \\
& \tau_a^\omega \in \{0, 1\}, \quad a \in E(i), i \in N, \omega \in W
\end{aligned} \tag{44}$$

The constraints in previous model M0 have been derived in previous subsections and will be explained briefly here. Constraints A1 and A2 (see formulas (11) and (24)) relate the availability of bus units, their assignment to bus lines and the number of services that the fleet assigned to a line must perform on that line. Constraints A3 and A4 are derived in Sect. 4.1.4 (see formula (29)), and their purpose is that, if one or more bus units are assigned to a line, then frequencies on that line must be within the interval $[H/h_{\max}, \hat{f}]$. The feasibility of passenger flows according to demands on the expanded graph is expressed by constraint B1 (see formula (10)). Constraint Qb1, as derived in Sect. 4.1.3 (see formula (28)), is a limitation on the total input flow that a bus stop is able to admit in order not to (a) surpass the ratio of services that bus service time permits, (b) produce spillback according to the permitted space for bus queue lengths or (c) blocking between buses having finished passenger service and ready to depart from their berths and those buses queueing for entering to the external traffic flow. Constraint Qp1 expresses the limitation imposed by the overall line capacity $c z^\ell$ and passenger flow on that line at a station. Finally, constraint Qp2, as derived in Sect. 4.3 (see formula (4.3)), has the purpose of limiting the queue length of passengers waiting at a bus stop according to the space available there. The routing

constraints R1 and R2 enforce that only one exiting link from a node $i \in N$ in the expanded transit network can be used for routing passengers of an origin–destination pair $\omega \in W$. The constant M in constraint R2 is a number verifying $M \geq \sum_{\omega \in W} g_\omega$.

A closed form for functions $\zeta_a(v, z)$ providing the total passenger queueing time at bus stops is not known and their values have been obtained by simulation and by solving numerically bulk-service queueing models (see Appendix). In addition, the second summation in the objective function is not composed of convex and differentiable functions in (v, z) in the general case, and that the functions in constraints A2 are non-convex and non-differentiable as well. The same applies to the family of constraints Qb1, which actually splits into three families of constraints as expressed in (28) also resulting in non-convexities.

A remarkable feature of this model is that it enables the detection of bottlenecks that limit the capacity of the bridging system. This allows for the determination of which bus stops are operating at their maximum number of bus services, the maximum number of buses that will queue at a bus stop, their average queueing time and the maximum number of passengers who are expected to remain waiting at a bus stop. The model also evaluates the passenger service times required at bus stops for each of the lines operating on them, and the average passenger queueing time at bus stops based on a bulk-service queueing model that assumes a fixed coefficient of variation for the interarrival times of buses (headways).

A possible extension of model M0 for setting frequencies in metro lines is now briefly discussed. In this case, no queueing of units at stations would take place ($w_b^0(\cdot) \equiv 0$), no waiting system L1 exists ($w_b^0(\cdot) \equiv 0$) and constraint Qb1 would be reduced to a modified version of constraint (27), with $s_b = 1$. The efficiency factor $e(s)$ for stations would be 1 in all expressions. Other components of the model that would require modification include the passenger service time formulas in Sect. 4.1.1. Particularly, in this case, formula (15) for parallel access/egress of passengers should be substituted by serial formulas (see, for instance Guenther and Sinha 1983; Fernández and Tayler 2005; Jaiswal et al. 2008; Fernández 2011). In this context, however, the inclusion in the model of modal split relationships would be required.

Section 6 develops an approximation for functions $\zeta_a(v, z)$, i.e., the total waiting times at bus stops, which permits the use of the numerical solutions obtained for bulk-service queues.

6 Approximating total waiting time of passengers at bus stops

Model M0 is a mixed-integer nonlinear programming problem, formulated more as a working framework than as an operating mathematical model because no closed forms are known for functions $\zeta_a(v, z)$. Nonlinearities, both in continuous and in integer variables, make this model very difficult to solve. In addition, functions for queueing delays both for passengers and buses at stops are complex, and, in the case of queueing of passengers, there are not closed-form expressions as discussed in Appendix.

To devise a computationally feasible procedure for finding approximations to model M0, the waiting time $w_b^0(v, z)$ of buses at stops will be approximated using

Allen–Cunneen’s formula in Allen (1998) and an approximation of the total waiting time of passengers at bus stops, $\zeta_a(v, z)$, is required in closed form. An approximated formulation in closed form that is suitable for the purposes of dealing with these delays in a mathematical programming program is developed in the following subsection.

In Sect. 4.2 it has been shown that the total waiting time ζ_a spent by passengers at a queue $a = a(\ell, b)$ is of the form $\zeta_a(v, z) = v_a P_a(z) \xi_a(\varrho_a(v, z))$, where $P_a(z)$ is of the form (33), and function $\xi_a(\cdot)$ is defined in (30). The following assumption will be supposed to hold for function $\xi_a(\cdot)$. Because we will now focus on a generic boarding queue $a = a(\ell, b)$, for ease of reading, subscript a will be omitted in this section.

Assumption *Function $\xi(\cdot)$ is non-decreasing and convex in $[0, 1]$ and verifies (a) $\xi(0) = 1$ and (b) $\lim_{\varrho \rightarrow 1} \xi(\varrho) = \infty$.*

We will focus our attention on function $\phi(v_a, v_x)$, which has as arguments boarding flow v_a and on-the-bus flow v_x .

$$\phi(v_a, v_x) = v_a \xi\left(\frac{v_a}{\eta - v_x}\right) \quad (45)$$

Clearly, ϕ is the mean overall waiting time for passengers willing to board on buses of a given line at a stop per period of service. Under these conditions, it is shown in the next proposition that function ϕ is convex in (v_a, v_x) and that, accordingly, it can be approximated by a finite set of cutting planes to the surface $(\phi(v_a, v_x), v_a, v_x)$ in \mathbb{R}^3 ; thus, it can easily be incorporated into an optimization model.

Proposition *If function $\xi(\cdot)$ verifies previous assumption then, function ϕ is convex on the set $B = \{(v_a, v_x) \in \mathbb{R}^2 \mid v_a + v_x \leq \eta, v_a \geq 0, v_x \geq 0\}$.*

Proof ξ' will stand for $\frac{d\xi}{d\varrho}$, and ξ'' will stand for $\frac{d^2\xi}{d\varrho^2}$. The Hessian of function ϕ in (45) is then

$$\begin{pmatrix} \frac{\partial^2 \phi}{\partial v_a \partial v_x} \end{pmatrix} = \frac{\varrho \xi'' + 2\xi'}{\eta - v_x} \begin{pmatrix} 1 & \varrho \\ \varrho & \varrho^2 \end{pmatrix} \quad (46)$$

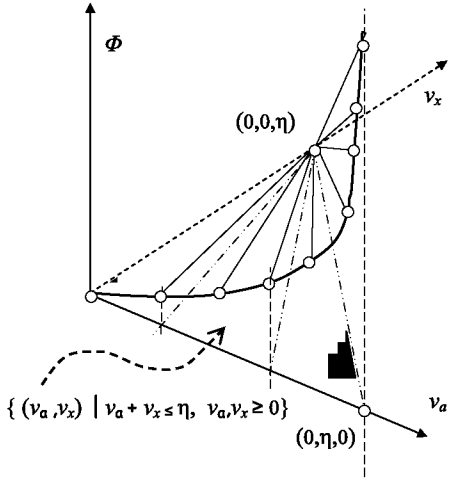
with eigenvalues $\lambda_1 = 0, \lambda_2 = (1 + \varrho^2)(\varrho \xi'' + 2\xi')(\eta - v_x)^{-1}$. Thus, the previous Hessian is positive semidefinite on B (or equivalently $\Rightarrow \varrho \leq 1$) because $\varrho \xi'' + 2\xi' \geq 0$ by the convexity of ξ . \square

In addition to the previous proposition, we can easily check that the surface $(\phi(v_a, v_x), v_a, v_x) \in \mathbb{R}^3$ can be obtained using straight segments leaning on the point $(0, 0, \eta) \in \mathbb{R}^3$ and sliding on the curve $(v_a \xi(\frac{v_a}{\eta}), v_a, 0)$, as shown in Fig. 4.

An approximation can be made to ϕ on $B = \{(v_a, v_x) \in \mathbb{R}^2 \mid v_a + v_x \leq \eta, v_a \geq 0, v_x \geq 0\} (\Rightarrow \varrho \leq 1)$ using its values and gradients at a set S of n_ϕ points:

$$S = \{(v_a^{(k)}, 0) \in \mathbb{R}^2 \mid v_a^{(k)} \in [0, \eta], 0 \leq k \leq n_\phi - 1\} \quad (47)$$

Fig. 4 Representing total delay per service period ϕ as a function of flows (v_a, v_x)



If $\phi_a^{(k)} = \phi(v_a^{(k)}, 0)$, $\hat{\phi}_a^{(k)} = \nabla_{v_a} \phi(v_a^{(k)}, 0)$, $\hat{\phi}_x^{(k)} = \nabla_{v_x} \phi(v_a^{(k)}, 0)$ with $(v_a^{(k)}, 0) \in S$, then an approximation of ϕ on B can be made with the convex function ψ :

$$\psi(v_a, v_x) = \max_{0 \leq k \leq n_\phi - 1} \{ \phi_a^{(k)} + \hat{\phi}_a^{(k)}(v_a - v_a^{(k)}) + \hat{\phi}_x^{(k)} v_x \} \quad (48)$$

more explicitly

$$\psi(v_a, v_x) = \max_{0 \leq k \leq n_\phi - 1} \left\{ v_a^{(k)} \xi(\varrho_k) + (\xi(\varrho_k) + \varrho_k \xi'(\varrho_k)) (v_a - v_a^{(k)} + \varrho_k^2 \xi'(\varrho_k) v_x) \right\} \quad (49)$$

where $\varrho_k = \frac{v_a^{(k)}}{\eta}$. If $v_a^{(k)} = \eta k / (n_\phi - 1)$, $k = 0, 1, 2, \dots, n_\phi - 1$, then it follows that $\varrho_k = k(n_\phi - 1)^{-1}$, $k = 0, 1, 2, \dots, n_\phi - 1$.

6.1 Application of this approximation to model M0

In the model, the total alighting capacity for a number of z services will be given by $\eta = cz^\ell$, where c is the average maximum capacity of a bus, and z^ℓ is the number of services received at the bus stop during the period of h minutes. For a link $a = a(\ell, b)$, $b \in L_b$, $\ell \in L$

$$\begin{aligned}
& v_a^{(k)} \xi(q_k) + (\xi(q_k) + q_k \xi'(q_k))(v_a - v_a^{(k)}) + \varrho_k^2 \xi'(q_k) v_x \\
& = -q_k \xi'(q_k) v_a^{(k)} + (\xi(q_k) + q_k \xi'(q_k)) v_a + \varrho_k^2 \xi'(q_k) v_x \\
& = -\varrho_k^2 \xi'(q_k) c z^\ell + \varrho_k^2 \xi'(q_k) v_x + (\xi(q_k) + q_k \xi'(q_k)) v_a \\
& = -c \tilde{\beta}_k z^\ell + \tilde{\beta}_k v_x + \tilde{\gamma}_k v_a \quad (0 \leq k \leq n_\phi - 1)
\end{aligned} \tag{50}$$

Thus, $\zeta_a(v, z)$, as defined in (35), can be approximated as follows:

$$\begin{aligned}\zeta_a(v, z) &= P_a(z^\ell)\phi(v_a, v_x) \approx P_a(z^\ell)\psi(v_a, v_x) \\ &= P_a(z^\ell) \max_{0 \leq k \leq n_\phi - 1} \{-c\tilde{\beta}_k z^\ell + \tilde{\beta}_k v_x + \tilde{\gamma}_k v_a\}\end{aligned}\quad (51)$$

Accordingly, if a new variable representing the total waiting time ζ_a of passengers queueing at a stop b for line ℓ is used, then the following constraints can be included in the mathematical programming formulation of design model M0:

$$P_a(z^\ell)(-c\tilde{\beta}_k z^\ell + \tilde{\beta}_k v_x + \tilde{\gamma}_k v_a) \leq \zeta_a, \quad 0 \leq k \leq n_\phi - 1 \quad (52)$$

or, equivalently,

$$\begin{aligned} \tilde{\beta}_{a,k} v_{x(a)} + \tilde{\gamma}_{a,k} v_a &\leq \frac{\zeta_a}{P_a(z^\ell)} + c^\ell \tilde{\beta}_{a,k} z^\ell, \\ a &= a(\ell, b), b \in \Pi_\ell, \ell \in L, 0 \leq k \leq n_\phi - 1 \end{aligned} \quad (53)$$

Coefficients $\tilde{\beta}_{a,k}, \tilde{\gamma}_{a,k}$ in the previous constraints are then

$$\begin{aligned} \tilde{\beta}_{a,k} &= \varrho_{a,k}^2 \xi'(\varrho_{a,k}), \\ \tilde{\gamma}_{a,k} &= \xi(\varrho_{a,k}) + \varrho_{a,k} \xi'(\varrho_{a,k}) \\ a &= a(\ell, b), b \in \Pi_\ell, \ell \in L, 0 \leq k \leq n_\phi - 1 \end{aligned} \quad (54)$$

7 An approximated model M1 and a heuristic to solve it

Modeling passenger waiting times at stops $b \in \hat{N}_G$ with the approximation provided by constraints (53) and using variables $\zeta_{a(\ell,b)}$, model M0 can be approximated by the following model M1:

$$\begin{aligned} [\text{Model M1}] \quad & \text{Min}_{n,z,v,\tau,\lambda,\xi} \sum_{\ell \in L} (\varsigma^\ell n^\ell + \gamma^\ell z^\ell) + \theta \sum_{a \in A} v_a T_a(v, z) + \theta \sum_{\ell \in L} \sum_{b \in \Pi_\ell} \zeta_{a(\ell,b)} \\ \text{s.t.:} \quad & \text{constraints A1 to A4, B1, R1, R2, Qb1 in model M0} \\ \text{Qp1} \quad & v_a + v_{x(a)} \leq c z^\ell \\ \text{Qp2} \quad & \sum_{\ell \in L_b} \zeta_a \leq \frac{H}{\eta_b} \hat{N}_b^{\text{pax}} \\ \text{Qp3} \quad & \tilde{\beta}_{a,k} v_{x(a)} + \tilde{\gamma}_{a,k} v_a \leq \frac{\zeta_a}{P_a(z^\ell)} + c^\ell \tilde{\beta}_{a,k} z^\ell, \quad 0 \leq k \leq n_\phi - 1 \\ \text{Qp4} \quad & \zeta_a \geq 0 \\ & a = a(\ell, b), b \in \Pi_\ell, \ell \in L \text{ in Qp1 to Qp4} \end{aligned} \quad (55)$$

In addition to the nonlinearities and non-convexities of model M0 described in Sect. 5, the approximation of functions $\zeta_a(v, z)$ using constraints Qp3 (also non-linear and non-convex) and Qp4 and variables ζ_a , add further difficulty in finding global minimal solutions. Thus, to find a computationally workable method for finding solutions for model M1, a specific heuristic method has been developed. To describe this heuristic method, model M2, below, must be considered. This model is a mixed-integer linear programming problem, that comes from partially freezing the nonlinearities of model M1 and setting proper values for functions $C_\ell(\cdot, \cdot), T_a(\cdot, \cdot)$,

$P_a(\cdot)$, $\hat{Z}_b(\cdot, \cdot)$, which provide, respectively, the line cycle lengths, the travel times on links, the uncongested average waiting time per passenger and per service and the maximum number of services that a bus stop is able to dispatch.

With next model, M2, the algorithm described below searches heuristically for a solution to model M1 using a fixed-point iteration based on the method of successive averages (MSA). In the algorithm below, function $\tilde{Z}_b(\cdot, \cdot)$ used in (58) is defined by

$$\tilde{Z}_b(\alpha, \beta) \triangleq H \min \left\{ \frac{\mathcal{L}^0}{\eta^0(\alpha + \beta)}, \frac{\mathcal{L}^1}{\eta^1 w_b^1} \right\}, \quad b \in N_G \quad (56)$$

$$\begin{aligned} \left[\begin{array}{c} \text{Model} \\ \text{M2}(\bar{T}, \bar{Z}, \bar{P}, \bar{C}) \end{array} \right] & \quad \text{Min} \quad \sum_{\ell \in L} (\varsigma^\ell n^\ell + \gamma^\ell z^\ell) + \theta \sum_{a \in A} v_a \bar{T}_a + \theta \sum_{\ell \in L} \sum_{b \in \Pi_\ell} \zeta_{a(\ell, b)} \\ \text{s.t.:} & \quad \text{constraints A1, A3, A4, B1, R1, R2 in model M0} \\ \text{A2} & \quad H^\ell n^\ell \geq z^\ell \bar{C}_\ell, \quad \ell \in L \\ \text{Qb1} & \quad \sum_{\ell \in L_b} z^\ell \leq \bar{Z}_b, \quad b \in \hat{N}_G \\ \text{Qp1} & \quad v_a + v_{x(a)} \leq c z^\ell \\ \text{Qp2} & \quad \sum_{\ell \in L_b} \zeta_a \leq \frac{H}{\eta_b} \hat{N}_b^{\text{pax}} \\ \text{Qp3'} & \quad \tilde{\beta}_{a,k} v_{x(a)} + \tilde{\gamma}_{a,k} v_a \leq \frac{\zeta_a}{\bar{P}_a} + c^\ell \tilde{\beta}_{a,k} z^\ell, \quad 0 \leq k \leq n_\phi - 1 \\ \text{Qp4} & \quad \zeta_a \geq 0 \\ & \quad a = a(\ell, b), b \in \Pi_\ell, \ell \in L \\ & \quad \text{in Qp1, Qp2, Qp3' and Qp4} \end{aligned} \quad (57)$$

Heuristic algorithm for model M1

0. (a) Determine initially suitable values for the number of services and an initial value for the uncongested waiting time per passenger and service $\bar{P}_a^{(0)}$, at a bus stop. Also set the initial line cycle lengths $\bar{C}_\ell^{(0)}$ for each line, the default bus service time at stops $\kappa_b^{(0)}$, and the initial values for bus queueing times at stops $w_b^{0,(0)}$, $b \in N_G$, so that an initial value for the maximum number of services allowable at a stop $\bar{Z}_b^{(0)}$ can be evaluated using function $\tilde{Z}_b(\cdot, \cdot)$ defined in (56), i.e., $\bar{Z}_b^{(0)} = \tilde{Z}_b(\kappa_b^{(0)}, w_b^{0,(0)})$. In addition, determine suitable link travel times $\bar{T}_a^{(0)}$ accordingly.
- (b) Solve model M2 for parameters $(\bar{T}^{(0)}, \bar{Z}^{(0)}, \bar{P}^{(0)}, \bar{C}^{(0)})$ to obtain flows and the number of services $(v^{(1)}, z^{(1)})$. Set $v = 0$

At iteration $v + 1$:

1. Obtain new values for the bus service time $\kappa_b^{(v+1)}$, the queueing time of buses at stops, $w_b^{0,(v+1)}$, and the maximum number of services allowed at each stop $\bar{Z}_b^{(v+1)}$

by an MSA step using $\alpha_v = 1/(v + 2)$:

$$\begin{aligned}
 \kappa_b^{(v+1)} &= \kappa_b^{(v)} + \alpha_v (\kappa_b(v^{(v+1)}, z^{(v+1)}) - \kappa_b^{(v)}) \\
 w_b^{0,(v+1)} &= w_b^{0,(v)} + \alpha_v (w_b^0(v^{(v+1)}, z^{(v+1)}) - w_b^{0,(v)}) \\
 \bar{Z}_b^{(v+1)} &= \bar{Z}_b^{(v)} + \alpha_v (\bar{Z}_b(\kappa_b^{(v+1)}, w_b^{0,(v+1)}) - \bar{Z}_b^{(v)}) \\
 b &\in N_G
 \end{aligned} \tag{58}$$

Update line cycles $\bar{C}_\ell^{(v+1)} = C_\ell(v^{(v+1)}, z^{(v+1)})$, $\ell \in L$, uncongested waiting time per passenger and per service at stop, $\bar{P}_a^{(v+1)} = P_a(z^{(v+1)})$, and link travel times $\bar{T}^{(v+1)}$ as follows:

$$\begin{cases} \bar{T}_a^{(v+1)} = t_a^0 + w_b^1 + \kappa_{b,\ell}(v^{(v+1)}, z^{(v+1)}) + w_{b'}^{0,(v+1)}, & \text{if } a = (j_\ell(b), j'_\ell(b')), \\ & \ell \in L, b, b' \in \Pi_\ell \\ \bar{T}_a^{(v+1)} = T_a(v^{(v+1)}, z^{(v+1)}) & \text{otherwise} \end{cases} \tag{59}$$

2. Solve model M2 for parameters $(\bar{T}^{(v+1)}, \bar{Z}^{(v+1)}, \bar{P}^{(v+1)}, \bar{C}^{(v+1)})$ and obtain flows $v^{(v+2)}$ and number of services $z^{(v+2)}$. Let $v \leftarrow v + 1$ and return to step 1.

The algorithm stops when, at a predetermined number r of consecutive iterations, the number of services assigned to bus lines does not change ($z^{(v+1)} = \dots = z^{(v+r)}$) and when, during these r iterations, flows v and total delays ζ have little fluctuation ($\|v^{(v+s+1)} - v^{(v+s)}\|_2 \leq \epsilon_v$ and $\|\zeta^{(v+s+1)} - \zeta^{(v+s)}\|_2 \leq \epsilon_\zeta$, $s = 1, 2, \dots, r - 1$). In the computational tests shown in Sect. 8, a restart of the MSA step is made every five or ten iterations.

8 Computational tests

To illustrate the model's performance and the proposed heuristic method for solving it, two test cases have been used.

The first case is a set of 20 candidate bus lines operating on six bus stops for a bridging system to assist after the disruption of a railway corridor in Madrid, composed of four railway stations. The expanded bus network, as defined in Sect. 3, consists of 118 nodes and 240 links, four centroids that correspond with the physical location of disrupted railway stations and 12 origin–destination pairs. The origin–destination trip table, with more than 37.000 passengers in a three-hour peak period, for Madrid's corridor, appears in Table 1 and was provided by railway authorities. The second test network is a set of 48 candidate bus lines operating on 17 bus stops for a bus-bridging system assisting after a disruption of the line 1 metro network in Barcelona (from stations Plaça d'Espanya to Clot). This expanded bus network consists of 310 nodes and 640 links, 10 centroids (corresponding to the disrupted metro stations) and 88 origin–destination pairs. The origin–destination trip table for Barcelona's line 1 has been estimated at a total of 37.992 trips, also over a period of three hours, and it is not shown for the sake of brevity. Figures 5 and 6 show, in the upper part, a schema of the streets on which the bus-bridging systems will operate

Table 1 The origin–destination trip matrix (station-to-station) during period $H = 180$ minutes for Madrid’s railway corridor. The *last row* and *column* are the average rates of arrivals and departures per minute at stations

| | At | Re | NM | Ch | Total Or. | pax/min |
|-------------|-------|-------|--------|------|---------------|---------|
| At | 0 | 2.011 | 22.097 | 368 | 24.476 | 135.98 |
| Re | 170 | 0 | 3.066 | 230 | 3.466 | 19.25 |
| NM | 4.386 | 150 | 0 | 170 | 4.706 | 26.14 |
| Ch | 2.504 | 150 | 2.438 | 0 | 5.092 | 28.28 |
| Total Dest. | 7.060 | 2.311 | 27.601 | 768 | 37.740 | – |
| pax/min | 39.22 | 12.84 | 153.34 | 4.26 | – | – |

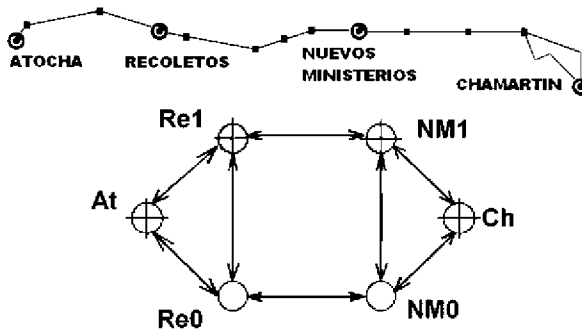


Fig. 5 (Top) A schematic representation of Paseo de la Castellana’s layout in Madrid where the bus-bridging system must operate during a representation of the four disrupted railway stations. (Bottom) A representation of the subgraph (N_G, A_G) for movements of passengers outside the bus-bridging lines. Bus stops at Recoletos and Nuevos Ministerios have been mirrored on both sides of the Paseo de la Castellana. Disrupted railway stations appear as nodes marked with a *cross*, which in this model are also considered as possible bus stops

in both test cases. Figure 5 depicts a schematic representation of Paseo de la Castellana’s layout in Madrid, where the bus-bridging system is intended to operate, and Fig. 6 shows a schema of the main arterials where the bus-bridging system may operate in Barcelona (Gran Via de les Corts Catalanes, Av. Meridiana, as well as some minor streets). Both expanded bus networks contain a set of links A_G (as defined in (1)) for the movement of passengers carried out outside the bus network (access from disrupted stations to bus stops, transfers between bus lines, portions of the trip carried out on foot). This set of links A_G is depicted at the bottom of Figs. 5 and 6. In both test networks, (a) bus units with a capacity of 100 passengers has been assumed; (b) the time period under consideration is a morning peak period of $H = 180$ minutes; (c) the value of time θ has been assumed to be 0.081 €/min ; and (d) the distances between bus stops have been realistically estimated, and it has been assumed that appropriate bus lanes exist to keep buses running at an average speed of 26 km/h between bus stops.

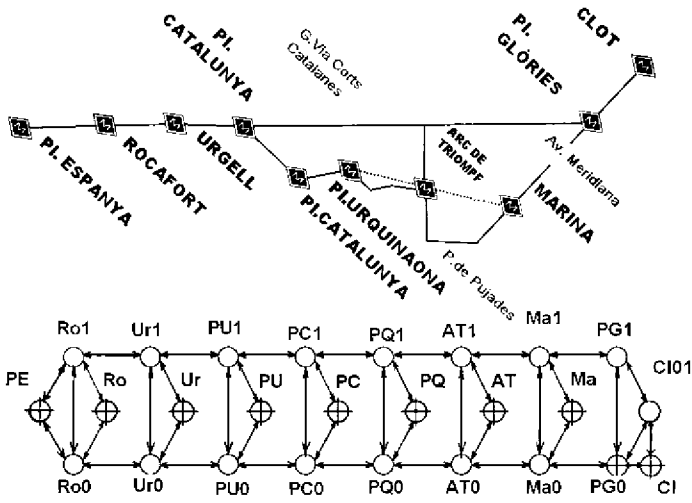


Fig. 6 (Top) A schematic representation of main arterials in Barcelona where the bus-bridging system must operate during a disruption of metro line 1. (Bottom) A representation of the subgraph (N_G, A_G) for the movements of passengers outside the bus-bridging lines. Excluding Plaça d'Espanya and Clot, all other bus stations have been mirrored on both sides of the arterials. Disrupted metro stations appear as nodes marked with a cross

Table 2 The results for the auxiliary bus system for Barcelona's metro line 1 (a) when all passengers for an origin–destination pair are forced to follow a single route and (b) when passengers of an origin–destination pair are allowed to board on the auxiliary system at different bus stops, choosing there a unique bus line, which may depend on the passenger's origin–destination pair

| Run # | η | Dem trips. | UnCov. | Tcpu (s) | #iter | AvT (min) | % Walk | \bar{v} (km/h) | f^* (€) | $f^{(0)}$ (€) | |
|-------|--------|------------|--------|----------|-------|-----------|--------|------------------|------------|----------------|---------|
| 1 | a | 0.5 | 18.996 | 29% | 929 | 140 | 27 | 56% | 6.9 | 41.552 | 74.213 |
| 2 | a | 1.0 | 37.992 | 52% | 1.958 | 151 | 34.9 | 76% | 5.5 | 107.290 | 135.380 |
| 3 | a | 1.3 | 49.130 | 67% | 3.366 | 251 | 37.3 | 83.4% | 5.1 | 149.190 | 175.230 |
| 4 | b | 0.5 | 18.996 | 22.7% | 797 | 150 | 27 | 56.2% | 6.9 | 41.546 | 74.213 |
| 5 | b | 1.0 | 37.992 | 46.5% | 1.144 | 180 | 34.8 | 73.4% | 5.4 | 106.840 | 135.380 |
| 6 | b | 1.3 | 49.130 | 65% | 681 | 140 | 37.1 | 82.4% | 5.1 | 148.420 | 180.280 |
| 7 | c | 1.3 | 49.130 | 51.5% | 183 | 31 | 35.2 | 76.5% | 7.2 | 140.880 | 180.680 |

The runs shown in this section for both test cases assume that the exploitation and setting costs of the bus-bridging system are nearly zero; thus, the objective of minimization is the total travel time of passengers.

Tables 2 and 3 show the main results. The heuristic algorithm has been implemented in AMPL, using CPLEX 11.0 as the solver. Runs shown in the previously mentioned tables have been performed on a 1.2 GHz laptop, 2 Gb RAM. To determine how the increase in the level of demand congests the bus network and affects the performance of the heuristic algorithm, several runs have been performed with the original origin–destination trip matrix shown in Table 1, with each cell multiplied

Table 3 The results for the auxiliary bus system for Madrid’s railway corridor (**a**) when all passengers for an origin–destination pair are forced to follow a single route and (**b**) when passengers of an origin–destination pair are allowed to board on the auxiliary system at different bus stops, choosing there a unique bus line, which may depend on the passenger’s origin–destination pair

| Run # | | η | Dem trips. | UnCov. | Tcpu (s) | #iter | AvT (min) | % Walk | \bar{v} (km/h) | f^* (€) | $f^{(0)}$ (€) |
|-------|---|--------|------------|--------|----------|-------|-----------|--------|------------------|-----------|---------------|
| 1 | a | 0.5 | 18.870 | 14.3% | ~6.8 | 51 | 25.3 | 40% | 12 | 38.741 | 70.521 |
| 2 | a | 1.0 | 37.740 | 72.8% | ~8.5 | 37 | 54.3 | 92.1% | 7.1 | 181.140 | 181.230 |
| 3 | a | 1.3 | 49.062 | 72.8% | ~8.6 | 42 | 59.5 | 91.8% | 5.5 | 236.270 | 238.250 |
| 4 | b | 0.5 | 18.870 | 14.3% | ~5.5 | 14 | 25 | 40% | 12 | 38.741 | 70.521 |
| 5 | b | 1.0 | 37.740 | 14.3% | ~9 | 21 | 41.7 | 63.8% | 7.8 | 127.350 | 168.520 |
| 6 | b | 1.3 | 49.062 | 14.3% | ~9.5 | 23 | 46.3 | 71.7% | 5.5 | 184.170 | 239.360 |

by a common factor η . Thus, $\eta = 1.3$ implies that all cells of the matrix have been increased by 30%, and $\eta = 1$ implies that the original trip table is being used. Column Dem shows the total resulting trips. Column UnCov reports the percentage of the trips that have not been performed on the auxiliary bus system. Column Tcpu reports the total number of elapsed seconds for the run and column #iter, the total number of iterations required by the heuristic algorithm described in Sect. 7 to converge. Column AvT reports the average trip time in minutes for the mass of passengers, and column \bar{v} reports their average speed. This speed is calculated as the total time required to go from origin to destination divided by the shortest distance that travelers would use by walking. Column %Walk reports the percentage of the total time used by passengers to reach their corresponding bus stops, transfer between bus stops or walk all the way from origin to destination. Finally, columns f^* and $f^{(0)}$ report the final objective function values in € reached by the heuristic method and the initial objective function value at the starting solution, respectively.

In both tables, runs marked with ‘a’ correspond to solutions of the model M0 where a single route is imposed on passengers of a given origin–destination pair on the expanded transit network, as stated in constraints R1 and R2. The results show that, as the level of demand increases, these routes become more and more congested, making travel by walking more attractive for a given origin–destination pair. This occurs particularly with those origin–destination pairs with a large number of trips. For runs marked with ‘b’ and also marked with ‘c’ in Table 2, passengers of a given origin–destination pair may ride a bus at different bus stops, but at any bus stop that they reach, they wait for a determined bus line without following a behavior based on strategies. In this last case, the performance of the bus-bridging system is slightly enhanced.

In all of the runs shown, the bottlenecks appear in the space of bus stops allotted for bus queues, which impose by means of constraints Qb1 in model M0, a limitation on the input flow of buses that the bus stop may admit. In the auxiliary bus system for Madrid’s corridor, forcing a unique route for all passengers of an origin–destination pair or allowing them to choose more than one bus stop substantially affects the performance of the solutions, because of the big flow from Atocha (At) to Nuevos Min-

isterios (NM), which cannot be allocated by a single bus line during the period of 3 hours for $\eta = 1.0$ and $\eta = 1.3$.

In the auxiliary bus system for Barcelona's metro (line 1), runs 3 and 6 (Table 2) performed with an extra 30% of demand. Under these conditions, bus stops at Plaça d'Espanya (PE) and Arc de Triomf (AT1) supported the maximum input flow of buses that the available queueing space of these stations allow (a greater flow would result in overflowing of the queueing space).

To illustrate how the model is able to detect different types of bottleneck in the system, Tables 4 to 7 show the operational parameters provided by the model results for runs 6 and 7 of Barcelona's network. Run 7 corresponds to identical demand levels and routing conditions as compared to run 6, but for a configuration in which the queueing space at bus stations PE and AT1 has been increased by two extra units. For run 6, three out of the 48 candidate lines have been assigned a positive number of services. These lines appear listed in Table 4 (lines L11ES, L1S3, L5ES). For run 7, an additional line, L1ES, has been added to the set of lines already selected in run 6. In both runs, all lines are point-to-point lines, i.e., with no intermediate stops. In Tables 4 and 6, column P3(a), after column from (to), denotes the average queueing time of passengers until they ride a bus of line ℓ on that station (time P3(a), as described in Sect. 2). Columns Occ+ and Occ- denote the ratio of passenger flow to line passenger capacity (cap_ℓ) in the direction from to to and in the opposite direction, respectively. Column $\kappa_{b,\ell}^+$ ($\kappa_{b,\ell}^-$) denotes the specific bus service time for line ℓ at bus stop from (to), according to formula (16). A value for maneuvering times $\kappa_b^M = 0.56$ min has been adopted. Values in (15) for single door boarding $g^+(1) = 1.8$ s/pax and two-rear-door-alighting $g^-(2) = 0.5$ s/pax have been used in all the runs, taking into account 5 additional seconds for the closing/opening of doors. z^ℓ denotes the total number of services carried out by the n^ℓ vehicles assigned to line ℓ during the time horizon. C^ℓ denotes the cycle length, and head. denotes the average time between bus arrivals at a stop of the line, or line headway. In Tables 5 and 7, characteristics and operational resulting parameters are shown for stations of the selected bus lines. Columns s_b , L^0 and $e(s_b)$ show the number of boarding platforms, the maximum number of buses allowed in the station (in boarding platforms + queueing) and the efficiency factor for the station, respectively. Columns L0oc, Cap and flow denote the peak occupation of the station in number of bus units, the maximum input flow of buses per hour allowed by the station and the actual input flow of buses, respectively. Columns w_b^0 and κ_b denote the average queueing time of buses and the average bus service time at the station. Finally, column PaxOc denotes the peak occupation of waiting passengers at the station.

Analyzing Tables 4 and 5 for run 6, the capacity (Cap) of stations PE and AT1 was the source of a bottleneck that did not allow for additional bus lines. The capacity could not be higher because the peak occupancy L0oc was at its maximum level at these stations. All other stations still had the capacity to allocate a higher bus input flow. Although it is not reported in the tables, it must be also remarked that bus service time (composed of by times V3b1, V3c and V3b2 described in Sect. 2) has not appeared as a limitation for the input flow of buses that a bus stop may admit. In all runs, the number of passengers that a bus station may accommodate who are queueing for his/her bus to arrive has been shown to be sufficient. A default average

Table 4 The model results of run 6 for the auxiliary bus system for Barcelona, and the resulting bus line parameters for the selected bridging lines with a positive number of services

| Bus line ℓ | from | P3(a) (min) | $\kappa_{b,\ell}^+$ (min) | Occ+ (%) | to | P3(a) (min) | $\kappa_{b,\ell}^-$ (min) | Occ- (%) | z^ℓ # | n^ℓ # | C_ℓ (min) | head. (min) | cap_ℓ (p/h) |
|-----------------|------|----------------|------------------------------|-------------|------|----------------|------------------------------|-------------|---------------|---------------|-------------------|----------------|----------------------------|
| L11ES | PE | 5.89 | 3.85 | 75.83 | CI01 | 6.72 | 4.48 | 80.32 | 42 | 8 | 33.88 | 4.28 | 1400 |
| L1S3 | AT1 | 2.41 | 2.06 | 23.08 | Ma1 | 3.49 | 4.93 | 60.64 | 49 | 5 | 16.12 | 3.67 | 1633 |
| L5ES | PE | 4.63 | 3.65 | 70.82 | AT1 | 6.59 | 4.06 | 81.23 | 44 | 7 | 25.80 | 4.09 | 1467 |

Table 5 The model results of run 6 for the auxiliary bus system for Barcelona, and the bus stop characteristics and their resulting operational parameters as output of the model

| Bus Stop b | s_b # | \mathcal{L}^0 # | L0oc # | PaxOc # | $e(s_b)$ | Cap (bus/h) | flow (bus/h) | w_b^0 (min) | κ_b (min) |
|------------|------------|----------------------|-----------|------------|----------|----------------|-----------------|------------------|---------------------|
| AT1 | 5 | 6 | 5.99 | 438 | 0.76 | 45.1 | 45.0 | 0.35 | 2.31 |
| CI01 | 6 | 7 | 3.13 | 378 | 0.68 | 31.3 | 14.0 | ~ 0 | 4.48 |
| Ma1 | 8 | 9 | 4.81 | 173 | 0.50 | 56.8 | 30.3 | ~ 0 | 3.17 |
| PE | 5 | 6 | 5.95 | 554 | 0.76 | 28.9 | 28.7 | 0.40 | 3.74 |

Table 6 Model results of run 7 for the auxiliary bus system for Barcelona. Bus line resulting operational parameters for the selected bridging lines with positive number of services

| Bus line ℓ | from | P3(a) (min) | $\kappa_{b,\ell}^+$ (min) | Occ+ (%) | to | P3(a) (min) | $\kappa_{b,\ell}^-$ (min) | Occ- (%) | z^ℓ # | n^ℓ # | C_ℓ (min) | head. (min) | cap_ℓ (p/h) |
|-----------------|------|----------------|------------------------------|-------------|------|----------------|------------------------------|-------------|---------------|---------------|-------------------|----------------|----------------------------|
| L11ES | PE | 4.44 | 3.59 | 69.20 | CI01 | 6.51 | 4.52 | 81.23 | 46 | 9 | 33.93 | 3.91 | 1533 |
| L1S3 | AT1 | 2.83 | 2.45 | 40.50 | Ma1 | 3.14 | 4.93 | 60.64 | 62 | 6 | 16.79 | 2.90 | 2066 |
| L5ES | PE | 3.06 | 2.89 | 51.57 | AT1 | 6.35 | 4.06 | 81.23 | 48 | 7 | 25.58 | 3.75 | 1600 |
| L1ES | PE | 4.17 | 2.90 | 51.90 | PC0 | 6.33 | 3.42 | 71.90 | 31 | 4 | 22.59 | 5.80 | 1033 |

Table 7 Model results of run 7 for the auxiliary bus system for Barcelona. Bus stop characteristics and their resulting operational parameters as output of the model

| Bus Stop b | s_b # | \mathcal{L}^0 # | L0oc # | PaxOc # | $e(s_b)$ | Cap (bus/h) | flow (bus/h) | w_b^0 (min) | κ_b (min) |
|------------|------------|----------------------|-----------|------------|----------|----------------|-----------------|------------------|---------------------|
| AT1 | 5 | 8 | 7.96 | 513 | 0.76 | 52.3 | 52.0 | 0.62 | 2.44 |
| CI01 | 6 | 7 | 3.46 | 406 | 0.68 | 31.0 | 15.3 | ~ 0 | 4.52 |
| Ma1 | 8 | 9 | 5.95 | 197 | 0.50 | 54.5 | 36.0 | ~ 0 | 3.30 |
| PC0 | 4 | 5 | 2.84 | 235 | 0.82 | 73.3 | 41.7 | ~ 0 | 1.36 |
| PE | 5 | 8 | 7.96 | 474 | 0.76 | 41.9 | 41.7 | 0.67 | 3.15 |

value of 20 seconds has been assumed for staying time w_b^1 at space L1 with capacity for one bus unit at all bus stops. These values have been shown to be adequate to produce an average occupation of space L1 shorter than 1, and, thus, no blocking has appeared. When changing the configuration of stations PE and AT1 allowing them a greater capacity in run 7, a new line is able to operate and more services may be allocated to the previous lines. Other changes in the configurations of stations PE and AT1 leading to an increase of their capacity, such as, increasing the number of boarding platforms, would have produced similar results.

Comparing the results for runs 6 and 7, the main conclusions are as follows:

- (a) The increase in the number of queueing spaces at stations PE and AT1 has global effects on the whole system and not only on the local performance of these stations. This change in the configuration allows for a moderate enhancement of the system because it permits more passengers to be transported by bus with better average traveling times.
- (b) However, this is possible at the cost of increasing the following local magnitudes: the average queueing time of buses at bus stops (w_b^0) and bus service times (κ_b).

Although they are not reported here, similar comparisons altering the capacity of the bus stops have also been made in the Madrid test case, with analogous results and the same conclusions. Finally, the level of congestion has a serious impact on the performance of the heuristic algorithm on these test cases.

9 Conclusions and further research

In this paper, a general model for setting frequencies on a set of transit lines that must operate under congested conditions has been developed, and it has been adapted to be used for setting the operating frequencies of a predetermined set of bus lines bridging a set of disrupted stations of a rapid transit or metro system. The model can be considered static and takes into account the average passenger flows over a given period, and considers the main factors intervening in the congestion of bus lines. The model provides as basic output the number of bus units to be assigned to each candidate bus line and the operating frequencies of these units. It also enables the detection of bottlenecks that may be present in the operation of the selected lines in terms of which bus stops operate at maximum number of bus services, the maximum number of buses able to queue at a bus stop and the maximum number of passengers who can be accommodated waiting at the bus stop. The model also provides an estimation of the average queueing time of buses waiting during the boarding/alighting of passengers, passenger service times required at bus-stops for each of the lines operating on them and the average queueing time of passengers at bus stops waiting for bus arrivals based on a bulk-service-type queueing model that assumes a fixed coefficient of variation for the interarrival times of buses (headways). The model has been run on two realistic test cases: a bus-bridging system for a railway corridor in Madrid with four rail stations and a bus bridging system for a segment of a metro line in Barcelona made up of ten metro stations. In this paper, the performance of the model on these two test cases is illustrated under different levels of demand. In addition, for one of

the test cases, the identification of the bottleneck points of the bus-bridging system is completed, and a detailed comparison of the resulting operational parameters is reported under different configurations of the bus stations.

The model is formulated in mixed-integer nonlinear programming, and it is solved using an *ad hoc* developed heuristic method, that allows solutions to be obtained within reasonable computational times; the performance of this heuristic and the validity of the solutions is shown in the previously described test cases. With small modifications, the presented model can be considered for setting frequencies on transit systems other than bus-bridging lines. In addition, the computational times obtained permit consideration of the model for use in solving larger instances than those presented. The exploration of other algorithmic alternatives is being considered by the authors in a forthcoming paper.

The assignment of passengers to bus lines is assumed to follow a pattern based on a single route on the expanded transit network for each origin destination pair because this provides an easy way to provide information to passengers willing to use the bridging system. Because the level of congestion on these systems must be high, the future lines of research being considered by the authors include the possibility of incorporating within a bus-bridging model the use of a congested-strategy-based passenger transit assignment similar to that described in Cepeda et al. (2006).

Acknowledgements Research supported under Spanish Research Projects TRA2008-06782-C02-01/02 and PT-2007-003-08CCPP

Appendix: A simple comparison of bulk service type queues and $M/M/s$ queues for bus stops

In this appendix, discrepancies are shown between the application of $M/M/s$ queueing models and a bulk-service model evaluated by simulation that accurately reproduces the congestion effects on a bus stop served by a single bus line.

A steady-state queueing model $M/M/1$ for the average waiting time $\varphi(\rho)$ of passengers at the bus stop queue as a function of the loading factor ρ of the bus stop following TCRP Web Document 6 (1999), would be as follows:

$$\varphi_{M/M/1}(\rho) = \frac{E[h]}{2} \left(1 + \frac{\text{Var}[h]}{E^2[h]} \right) + \frac{\rho}{\mu(1-\rho)} \quad (60)$$

The first term in (60) is the average waiting time per passenger at the bus stop resulting from bus headway characteristics of dispersion reflected in the coefficient of variation $\text{Var}^{1/2}[h]/E[h]$. If only this term applied, then implicitly, all passengers present at the bus stop could ride on the arriving bus, and no passenger would wait for a second bus (or third bus, ...). The second term is the contribution to the average queueing time per passenger of those passengers that cannot be served by the first-arriving bus. This second term depends on the load factor ρ and also on the rate of service μ . Under an exponentially distributed headway, then the first term would reduce to $E[h]$, as $\text{Var}^{1/2}[h]/E[h] = 1$ and that $\mu = c/E[h]$, with c as the average passenger capacity of buses arriving at the bus stop. Assume now that c is not random

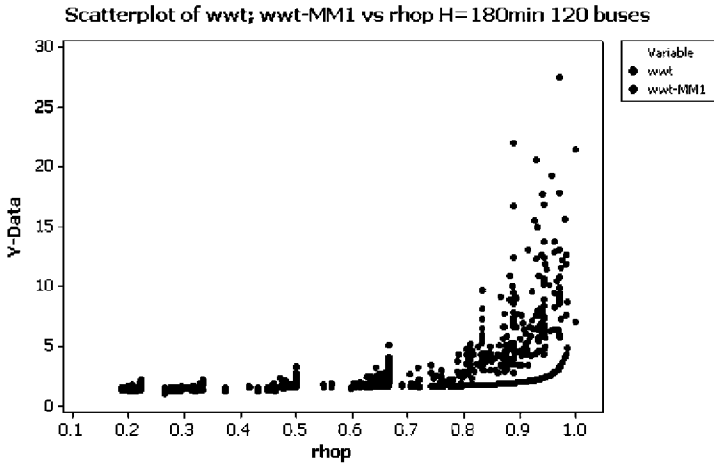


Fig. 7 A comparison between the average waiting times per passenger at the bus stop for bus arrivals with exponentially distributed headway for a period of 3 hours and the average waiting times for the $M/M/1$ queueing model. The capacity of arriving buses was assumed to be constant

and that $c = 30$ pax and that $E[h] = 90$ seconds so that the total alighting capacity of the line servicing the bus stop is 1200 pax/h.

To illustrate this comparison, a simulation model was developed in order to evaluate the average waiting time per passenger during a finite period of service H . The arrival of passengers at the stop was assumed to be Poissonian, and strict FIFO discipline of passengers was assumed to be observed. It was assumed that buses were empty on arrival and that the dwell time model at the stop was deterministic without holding. The probability distribution for headways could be chosen among several options. Figure 7 shows the simulation results for headways exponentially distributed and finite $H = 180$ minutes, (120 bus arrivals). These results appear as a scatterplot in which each point is a simulation shot. A point's x -coordinate is the resulting loading factor of the bus stop, and its y -coordinate is the waiting time in minutes. Function $\varphi_{M/M/1}(\rho)$ appears overlapped on the scatterplot, showing clearly that serious discrepancies of both models appear for $\rho > 0.6$, i.e., under high congestion, and that the $M/M/1$ model strongly underestimates passenger waiting times. In addition, the scatter of points presents a heavy dispersion as long as the load factor approaches 1. This dispersion is the result of the moderate but realistic horizon length H chosen.

Another set of simulation runs was made with strictly deterministic headways of arriving buses. In this case, the comparison with an $M/M/1$ model was made by applying formula (60) with $\text{Var}^{1/2}[h]/E[h] = 0$ but with same rate of service μ as before. (Notice that if $\varphi_{M/M/1}(\cdot)$ had been applied in this case under the form $\varphi_{M/M/1}(\rho) = 1/\mu + \rho/\mu(1 - \rho) = 1/\mu(1 - \rho)$, then $c/\mu > E[h]/2 =$ the uncongested waiting time at the bus stop.) The simulation results appear in Fig. 8 below showing that model $M/M/1$ overestimates the waiting times in this case. For a bus stop where buses arrive with random capacity, average waiting times per passengers would have been greater than those shown by the previous simulations. To evaluate the effect of randomness in passenger capacity of arriving buses, a steady-state

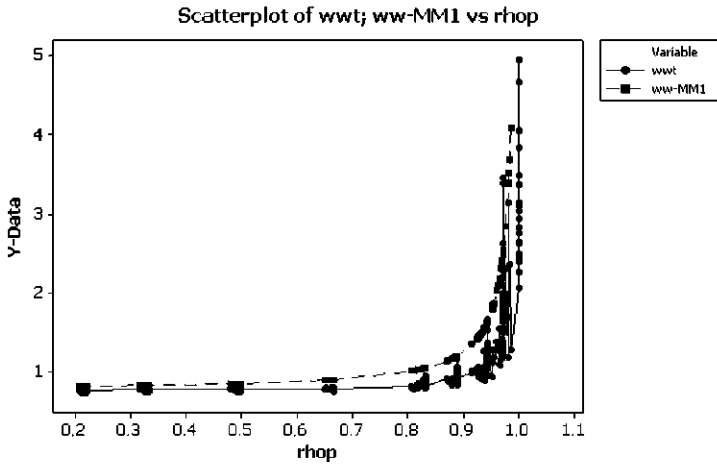


Fig. 8 A comparison between average waiting times per passenger at the bus stop for bus arrivals with strict constant headway for a period of 3 hours and the average waiting times for the $M/M/1$ queueing model

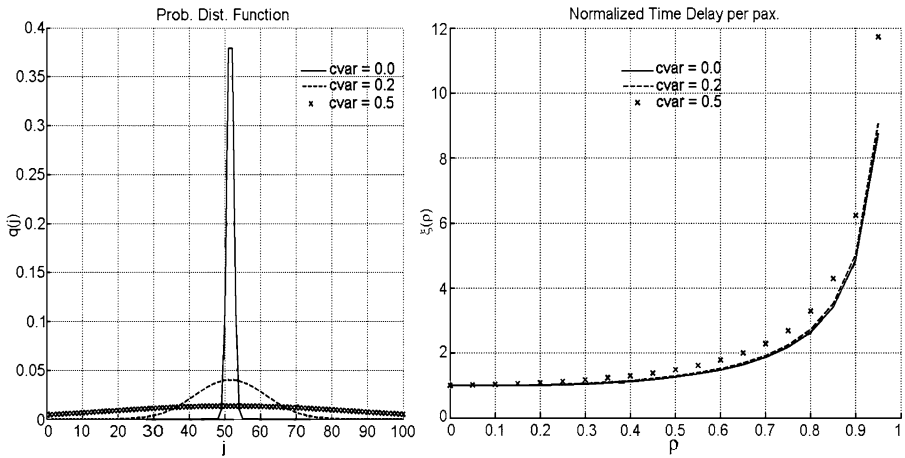


Fig. 9 The values of function $\xi(\rho)$ in (right) for a $M/M^{[Y]}/1$ bulk-service queue with random server capacity Y . Y is assumed to follow a truncated normal distribution on the interval $[0, 100]$ and with an average value of 50 pax for several coefficients of variation ranging from 0 to 0.5 (left)

bulk-service queue model $M/M^{[Y]}/1$ with random capacity was solved analytically for various probability distributions for the capacity. The effects of randomness in capacity of arriving buses on average waiting time per passenger appeared low or moderate, as shown in Fig. 9. This figure shows at right, the graph for function $\xi(\rho)$, i.e., the normalized delay per passenger and per service for three distributions of the capacity of arriving buses. All them are truncated normals on the interval $[0, 100]$ with an average of 50 passengers of capacity. The coefficients of variation for these distributions are 0.0, 0.2 and 0.5. For the $M/M^{[Y]}/1$ queueing model, Table 8, below,

Table 8 The coefficients $\tilde{\beta}$ and $\tilde{\gamma}$ in (54) for a bulk service queue $M/M^{[c]}/1$. The residual bus capacity c is distributed following a truncated normal on $[0, 100]$ with an average value of 50 pax. Values are reported for several coefficients of variation for c , C_c , in the range 0.0 to 0.5

| ρ | $C_c = 0.04$ | 0.10 | 0.2 | 0.3 | 0.4 | 0.5 |
|------------------|--------------|----------|----------|----------|----------|----------|
| $\tilde{\beta}$ | | | | | | |
| 0.05 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0003 | 0.0008 |
| 0.15 | 0.0015 | 0.0016 | 0.0026 | 0.0043 | 0.0073 | 0.0117 |
| 0.25 | 0.0225 | 0.0235 | 0.0270 | 0.0328 | 0.0405 | 0.0508 |
| 0.35 | 0.0952 | 0.0973 | 0.1043 | 0.1158 | 0.1312 | 0.1508 |
| 0.45 | 0.2709 | 0.2750 | 0.2880 | 0.3096 | 0.3388 | 0.3766 |
| 0.55 | 0.6694 | 0.6771 | 0.7024 | 0.7447 | 0.8036 | 0.8806 |
| 0.65 | 1.6263 | 1.6426 | 1.6972 | 1.7875 | 1.9156 | 2.0878 |
| 0.75 | 4.3380 | 4.3777 | 4.5128 | 4.7385 | 5.0655 | 5.5178 |
| 0.85 | 15.4343 | 15.5703 | 16.0344 | 16.8138 | 17.9800 | 19.6775 |
| 0.95 | 160.2669 | 161.6596 | 166.4096 | 174.5540 | 188.2396 | 212.0599 |
| $\tilde{\gamma}$ | | | | | | |
| 0.05 | 1.0000 | 1.0000 | 1.0000 | 1.0020 | 1.0110 | 1.0310 |
| 0.15 | 1.0110 | 1.0130 | 1.0200 | 1.0380 | 1.0740 | 1.1340 |
| 0.25 | 1.1110 | 1.1170 | 1.1370 | 1.1750 | 1.2350 | 1.3250 |
| 0.35 | 1.3490 | 1.3590 | 1.3910 | 1.4480 | 1.5330 | 1.6540 |
| 0.45 | 1.7830 | 1.7970 | 1.8450 | 1.9270 | 2.0470 | 2.2120 |
| 0.55 | 2.5720 | 2.5930 | 2.6650 | 2.7880 | 2.9670 | 3.2090 |
| 0.65 | 4.1500 | 4.1860 | 4.3060 | 4.5080 | 4.8010 | 5.2010 |
| 0.75 | 7.9790 | 8.0470 | 8.2810 | 8.6740 | 9.2480 | 10.0430 |
| 0.85 | 21.6390 | 21.8260 | 22.4650 | 23.5400 | 25.1460 | 27.4730 |
| 0.95 | 178.3520 | 179.8980 | 185.1750 | 194.2170 | 209.3520 | 235.5510 |

shows coefficients $\tilde{\beta}$ and $\tilde{\gamma}$ of the approximation (54) developed in Sect. 6.1 for several coefficients of variation of the capacity distributions for the arriving buses.

References

- Adebisi O (1986) A mathematical model for the headway variance of fixed-route buses. *Transp Res B* 1:59–70
- Allen AO (1998) Probability, statistics and queueing theory. Academic Press, New York
- Anderson PA, Scalia-Tomba GP (1981) A mathematical model of a bus route. *Transp Res B* 15(4):249–266
- Bailey NTJ (1954) On queueing process with bulk service. *J R Stat Soc B* 16:80–87
- Cepeda M, Cominetti R, Florian M (2006) A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. *Trans Res B* 40:437–459
- INRO Consultants, EMMÉ User's Manual. Software Release 9.1 2010
- Constantin I, Florian M (1995) Optimizing frequencies in a transit network: a non-linear bi-level programming approach. *Int Trans Oper Res* 2:149–164
- Fernández R, Tayler N (2005) Effect of passenger-bus-traffic interactions on bus stop operations. *Transp Plan Technol* 28(4):272–292

-
- Fernández R (2011) Modelling of bus-stop operations. In: Proceedings of the European transport conference. www.etcproceedings.org/paper/download/2637, last access June 21st (2011)
- Guenther RP, Sinha KC (1983) Modeling bus delays due to passenger boardings and alightings. *Transp Res Rec* 915:7–13
- Haupt T, Nökel K, Reiter U (2009) Schedule based dynamic transit assignment. In: Wilson NHM, Nuzzolo A (eds) *Schedule-based modeling of transportation networks*. Springer, Berlin
- Han J, Lee S, Kim J (2005) Meta-heuristic algorithms for a transit route design. In: *Advanced OR and AI methods in transportation*. <http://www.iasi.cnr.it/ewgt/16conference/ID26.pdf>, last access June 20th (2011)
- Jain JL, Mohanty SG, Böhm W (2007) *A course on queueing models*. Chapman & Hall, New York
- Jaiswal NK (1964) A bulk-service queueing problem with variable capacity. *J R Stat Soc B* 26:143–148
- Jaiswal SB, Jonathan M, Ferreira L (2008) Relating bus dwell time and platform crowding at a busway station. In: *Proceedings of the 1st Australasian transport research forum*, pp 239–240. Gold-Coast, Australia
- Kepaptsoglou K, Karlaftis MG (2008) Bus bridging problem in metro operations: mathematical framework and decision support system. In: *Transportation research board 87th meeting*. Washington, USA
- Noriega Y, Florian M (2003) L'optimisation des fréquences d'un réseau de transport en commun moyennement congestionné. *INFOR* 41(2):129–153
- Papola N, Filippi F, Gentile G, Meschini L (2009) Schedule based transit assignment: new dynamic equilibrium model with vehicle capacity constraints. In: Wilson NHM, Nuzzolo A (eds) *Schedule-based modeling of transportation networks*. Springer, Berlin
- Powell WB (1986) Approximate, closed form moment formulas for bulk arrival, bulk service queues. *Transp Sci* 20(1):13–23
- Spies H, Florian M (1989) Optimal strategies: a new assignment model for transit networks. *Transp Res, Part B, Methodol* 23(2):83–102
- Vasanthakumar NB (1993) Approximation for the variance of the waiting time in a $GI/G/1$ queue. *Microelectron Reliab* 33(13):1997–2002
- TCRP Web Document 6 (Project A-15). Transit capacity and quality of service manual. TRB 1999. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_webdoc_6-a.pdf, last access June 22nd 2011
- TCRP Report 19 (1996) Guidelines for the location and design of bus stops. Transportation research board national research council. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_19-a.pdf, last access June 21st 2011
- TCRP Report 100, 2nd edn (2003) Transit cooperative research program. Federal transit administration. TRB. <http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp100/part%200.pdf>
- TransCAD transportation planning software. Caliper Corporation. <http://www.caliper.com/tcovu.htm>, last access June 20th 2011
- Trivedi KS (2002) *Probability and statistics with reliability, queueing and computer science applications*, 2nd edn. Wiley, New York
- PTV AG | Global. PTV Vision VISUM 11.5. <http://www.ptvag.com/software/transportation-planning-traffic-engineering/software-system-solutions/visum/>, last access June 19th 2011
- Fan W., Machemehl RB (2004) Optimal transit route network design problem: algorithms, implementations, and numerical results. Research report SWUTC/04/167244-1. Southwest Region University Transportation Center for Transportation Research University of Texas at Austin
- Yoo G-S, Kim D-K, Chon KS (2010) Frequency design in urban transit networks with variable demand: model and algorithm. *KSCE J Civ Eng* 14(3):403–411
- Zhao F., Zeng X. (2008) Optimization of transit route network, vehicle headways and timetables for large-scale transit networks. *Eur J Oper Res* 186:841–855